PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 38/18, C12N 5/06, 5/08, A61K 35/34 — A1

(11) International Publication Number: **WO 98/27995**

(43) International Publication Date: 2 July 1998 (02.07.98)

(21) International Application Number: PCT/US97/23611

(22) International Filing Date: 19 December 1997 (19.12.97)

(30) Priority Data:
60/033,145 20 December 1996 (20.12.96) US

(71) Applicant *(for all designated States except US)*: CREATIVE BIOMOLECULES, INC. [US/US]; 45 South Street, Hopkinton, MA 01748 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: COHEN, Charles, M. [US/US]; 1 Harrington Lane, Weston, MA 02193 (US). SAMPATH, Kuber, T. [US/US]; 98 Pamela Drive, Hollison, MA 01746 (US).

(74) Agent: TWOMEY, Michael, J.; Testa, Hurwitz & Thibeault, LLP, High Street Tower, 125 High Street, Boston, MA 02110 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: TREATMENT OF MAMMALIAN MYOCARDIUM WITH MORPHOGEN LOCALLY, OR WITH MORPHOGENICALLY-TREATED MYOGENIC PRECURSOR CELLS

(57) Abstract

The present invention provides methods for the treatment, and pharmaceuticals for use in the treatment, of mammalian subjects at risk of, or afflicted with, loss of or damage to myocardial tissue. The methods involve the administration of certain morphogens, inducers of those morphogens, agonists of the corresponding morphogen receptors, or small molecule morphogenic activators, or implantation of cells induced with those agents. The morphogens useful in the invention include OP1, CBMP-2A (BMP-2), CBMP-2B (BMP-4), and other members of the morphogens family of the TGFβ superfamily of growth and differentiation factors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TREATMENT OF MAMMALIAN MYOCARDIUM WITH MORPHOGEN LOCALLY, OR WITH MORPHOGENICALLY-TREATED MYOGENIC PRECURSOR CELLS

Field of the Invention

The present invention relates generally to methods and preparations for the treatment of mammals, including humans, at risk of, or afflicted with, loss of or damage to myocardium. The methods involve the implantation of mammalian myogenic precursor cells treated with certain morphogens, inducers of those morphogens, agonists of the corresponding morphogen receptors, or with small molecule morphogenic activators.

Background of the Invention

Unlike skeletal muscle or smooth muscle, adult mammalian cardiac muscle has extremely limited powers of growth and regeneration. During development, the myocardium arises by end-to-end fusion of myogenic precursor cells to form branched myofibers in which individual cardiac myocytes are joined by intercalated disks. The myogenic precursor cells which give rise to the myocardium are derived from the splanchic mesoderm, which is derived from the lateral mesodermal mesenchyme which, in turn, arises from the mesoderm formed after gastrulation. It is generally believed that there are no remaining myogenic precursor cells in adult mammalian myocardium and, therefore, lost or damaged myocardium is typically replaced by fibrotic or scar tissue, rather than new myocardium. See, generally, B.M. Carlson, ed. (1981) Patten's Foundations of Embryology, 4th Edition, McGraw-Hill, New York. As a result, damage or loss of myocardium due, for example, to myocardial infarction, congestive heart failure, physical trauma (e.g., in an automobile accident), or infection, typically results in a permanent and often progressive loss of functional myocardium.

In contrast, mammalian skeletal muscle has much greater capacity for growth and regeneration, even in adulthood. Like the myocardium, skeletal muscle has its first origins after the induction of the mesoderm. After differentiation of the mesoderm into dorsal, intermediate, and lateral mesoderm, the dorsal mesodermal mesenchyme differentiates to form myotomes which, in turn, differentiate to form the myogenic precursor cells which ultimately form skeletal muscle. Unlike the myogenic precursor cells of the heart, the skeletal muscle precursors fuse side-to-side to form unbranched, multinucleated myofibers. Significantly, some portion of the

skeletal myogenic precursor cells do not differentiate into myocytes but, rather, attach to the plasmalemmas of the myocytes. These cells may remain, throughout adulthood, as largely undifferentiated, quiescent skeletal muscle "satellite cells." Upon injury of a skeletal muscle, however, these satellite cells are revealed to be myogenic precursor cells, or muscle "stem cells," which proliferate and differentiate into new and functional skeletal muscle. Even after injury, however, a portion of the proliferated satellite cells remain undifferentiated and attach to the newly formed myofibers. Thus, the satellite cells of skeletal muscle provide a constant and renewable source of myogenic precursor cells which allows for skeletal muscle repair and regeneration throughout mammalian life.

The proliferation and differentiation of skeletal muscle satellite cells has been extensively studied in vitro. For example, a simple saline extract of skeletal muscle has been shown to cause satellite cells to proliferate in culture (Bischoff (1989) in Myoblast Transfer Therapy, Griggs and Karpati, eds., pp. 147-158). Similarly, it has been shown that chick embryo extract or the conditioned medium of differentiated myotubes from young mice exhibits a strong mitogenic effect on satellite cells, but that conditioned medium from older murine myotubes has a lesser effect (Mezzogiorno et al. (1993) Mech. Ageing & Develop. 70:35-44). In addition, a number of hormones and growth factors have been found to enhance satellite cell proliferation, including FGF, PDGF, ACTH, LIF, and IGF (Bischoff (1989); Mezzogiorno et al. (1993)). Conversely, TGF-β_1 is widely believed to inhibit satellite cell proliferation, as does contact with the myofiber plasmalemma, but not the basal lamina (Bischoff (1989); but see Hathaway et al. (1991) J. Cell Physiol. 146:435-441).

Curiously, in a rat model of skeletal muscle injury, it was found that there were signs of satellite cell differentiation before there were significant signs of satellite cell proliferation (Rantanen et al. (1995) Lab. Invest. 72:341-347). This suggests the possibility that there are two populations of skeletal muscle satellite cells: "committed satellite cells" which respond to injury by rapidly differentiating to replace the injured tissue, and "stem satellite cells" which respond more slowly by proliferating and, perhaps, renewing the committed satellite cell population. In this scenario, the stem satellite cells may undergo mitosis to produce one daughter cell which remains a stem satellite cell, and another which becomes a committed satellite cell.

In another animal model, autologous mouse skeletal muscle cells were explanted from a healthy muscle, proliferated in vitro, and then implanted into a necrotized skeletal muscle site (Alameddine and Fardeau (1989) in Myoblast Transfer Therapy, Griggs and Karpati, eds., pp. 159-166). In these experiments, it was shown that the transplanted satellite cells were able to populate the necrotized area and differentiate into functional myotubes. Similarly, PCT Publication WO 96/28541 discloses that histocompatible donor mouse myoblasts can be implanted into the weakened muscle of a mouse model of muscular dystrophy and differentiate into myofibers. In addition, it is shown that growth of the myoblasts in bFGF results in significantly more new myofibers at the implant site. Thus, skeletal muscle satellite cells, proliferated in vitro, may be able to serve as a source of myogenic precursor cells for muscle restoration or regeneration therapy.

The ability of skeletal muscle satellite cells to restore or regenerate injured skeletal muscle, has led some researchers to test whether myogenic precursor cells could be used to replace lost or damaged myocardial muscle. For example, mouse fetal cardiomyocytes, which are not terminally differentiated and retain the ability to divide, have been directly injected into the myocardium of a syngeneic adult mouse, and have been shown to form new and apparently functional myocardium (Soonpaa et al. (1994) Science 264:98-101). Significantly, it has been shown that skeletal muscle satellite cells, explanted from adult canine skeletal muscle can be proliferated in vitro and implanted into a site of myocardial cryoinjury, where they appear to differentiate into "cardiac-like" muscle cells, possibly in response to morphogenic signals present in the myocardium (Chiu et al. (1995) Ann. Thorac. Surg. 60:12-18).

Morphogens and Growth Factors

A great many proteins have now been identified which appear to act as morphogenetic or growth factors, regulating cell proliferation and/or differentiation. Typically these growth factors exert their effects on specific subsets of cells and/or tissues. Thus, for example, epidermal growth factors, nerve growth factors, fibroblast growth factors, various hormones, and many other proteins inducing or inhibiting cell proliferation or differentiation have been identified and shown to affect some subset of cells or tissues.

One group of morphogenetic proteins, referred to herein as "morphogens," includes members of the family of bone morphogenetic proteins (BMPs) which were initially identified by

their ability to induce ectopic, endochondral bone morphogenesis. Subsequent characterization of the nucleic acid and amino acid sequences of the BMPs has shown them to be a subgroup of the TGFβ superfamily of growth and differentiation factors. Members of the morphogen family have now been shown to include the mammalian osteogenic protein1 (OP1, also known as BMP7), osteogenic protein2 (OP2), osteogenic protein3 (OP3), BMP2 (also known as BMP2A or CBMP2A), BMP3, BMP4 (also known as BMP2B or CBMP2B), BMP5, BMP6, Vgr1, and GDF1, as well as the Xenopus homologue Vgl and the Drosophila homologues DPP and 60A. Members of this family encode secreted polypeptides that share common structural features, and that are similarly processed from a pro-protein to yield a carboxy terminal mature protein of approximately 100-110 amino acids. All members share a conserved pattern of cysteines in this domain and the active form of these proteins is either a disulfide-bonded homodimer of a single family member, or a heterodimer of two different members (see, e.g., Massague (1990) Annu. Rev. Cell Biol. 6:597; Sampath, et al. (1990) J. Biol. Chem. 265:13198).

The members of the morphogen family of proteins are expressed naturally in a variety of tissues during development. BMP-2 (i.e., BMP-2A), for example, is expressed in embryonic mouse hair follicles, cartilage and bone (Lyons et al. (1989) Genes & Develop. 3:1657-1668); BMP3 has been shown to be most highly expressed in human embryonic lung and kidney, highly expressed in intestinal mucosa and skeletal tissues such as the perichondrium and periosteum, expressed in brain, but undetectable in embryonic heart and liver (Vukicevic et al. (1994) J. Histochem. Cytochem. 42:869-875); BMP4 has been shown to be expressed in the developing limbs, heart, facial processes and condensed mesenchyme associated with early whisker follicles in embryonic mice (Jones, et al. (1991) Development 111:531-542); and OP1 (BMP7) has been shown immunohistochemically to be present in human embryos in sclerotome, hypertrophied chondrocytes, osteoblasts, periosteum, adrenal cortex, renal convoluted tubules, placenta, smooth, cardiac and skeletal muscles, meninges and neural cells, as well as the basement membranes of the lungs, pancreas and skin (Vukicevic, et al. (1994) Biochem. Biophys. Res. Commun. 198:693-700). Some of the morphogens (e.g., OP2 and BMP2) were not detected in analyses of adult tissues, suggesting only an early developmental role for these morphogens (Ozkaynak, et al. (1992) J. Biol. Chem. 267:25220-25227).

Although, as noted above, several morphogens have been shown to be expressed in embryonic or adult mammalian heart tissue, and various utilities for the morphogens have been proposed and developed, it has never previously been shown or suggested that treatment of myogenic precursor cells with the morphogens, morphogen inducers, agonists of morphogen receptors, or small molecule morphogenic activators is useful in promoting the proliferation and/or differentiation of myogenic precursor cells into new and functional myocardium in a morphogenically permissive environment. Nor has it previously been shown or suggested that morphogenically-treated myogenic precursor cells are useful in the treatment of lost or damaged mammalian myocardium.

Summary of the Invention

The present invention is directed to methods of treatment, and pharmaceutical preparations for use in the treatment, of mammalian subjects at risk of, or afflicted with, loss of or damage to myocardium. Such subjects include subjects already afflicted with the loss of myocardial tissue, such as those which have already suffered a myocardial infarction, physical trauma to the heart (e.g., in an automobile accident, or those already suffering from congestive heart failure, as well as subjects reasonably expected to suffer from myocardial infarction or congestive heart failure. Whether a particular subject is at risk is a determination which may routinely be made by one of ordinary skill in the relevant medical or veterinary art.

In these methods of treatment, myogenic precursor cells are implanted into a mammal at a site at risk of, or afflicted with, loss of or damage to myocardium, and the myogenic precursor cells are morphogenically-treated prior to, simultaneously with, or subject to implantation. Thus, for example, morphogenically-treated mammalian myogenic precursor cells may be implanted into a mammalian heart at the site of a myocardial infarct, or into the damaged or weakened myocardium of a subject with congestive heart failure. The mammalian myogenic precursor cells may be derived from skeletal muscle (e.g., skeletal muscle satellite cells), from embryonic tissue (e.g., embryonic mesodermal mesenchyme) or from a myogenic precursor cell line maintained in vitro. Thus, the myogenic precursor cells may be derived from a donor (e.g., a tissue-type matched donor, sibling, identical twin, or fetus), may be derived from a tissue culture (e.g., undifferentiated or partly undifferentiated myogenic cells in culture; fetal tissue culture), or may be explanted from the subject and re-implanted after morphogen-induced proliferation and/or

differentiation. Finally, the morphogenic treatment of the implanted cells may include treatment of the cells with a morphogen, morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator prior to implantation, simultaneously with implantation, or subsequent to implantation.

The present invention is further directed to methods of promoting the proliferation and differentiation of mammalian myogenic precursor cells in vivo or in vitro. Thus, for example, myogenic precursor cells isolated from mammalian skeletal muscle tissue, embryonic myogenic precursor cells, or myogenic precursor cell lines, may be stimulated to proliferate by treatment with a morphogen, an inducer of a morphogen, an agonist of a morphogen receptor, or a small molecule morphogenic activator. Alternatively, or in addition, mammalian myogenic precursor cells may be stimulated to differentiate into myocytes, particularly myocytes which express markers of myocardial tissue, in a morphogenically permissive environment.

The present invention is further directed to therapeutic preparations comprising isolated mammalian myogenic precursor cells and an amount of a morphogen, inducer of a morphogen, agonist of a morphogen receptor, or small molecule morphogenic activator sufficient to promote proliferation or differentiation of the myogenic precursor cells in a morphogenically permissive environment.

The methods and compositions of the present invention capitalize in part upon the fact that certain proteins of eukaryotic origin, defined herein as morphogens, may be used to treat myogenic precursor cells such that, when these morphogenically-treated myogenic precursor cells are present in a morphogenically permissive environment, they may migrate, proliferate and/or differentiate so as to form new and functional myocardium. In particular, the present invention is based in part upon the fact that treatment of myogenic precursor cells with these morphogens enhances or increases the probability, rate, or efficiency with which these cells migrate, proliferate and/or differentiate into new and functional myocardium in a morphogenically permissive environment. Thus, in accordance with the present invention, morphogenically-treated myogenic precursor cells may be used to restore or regenerate lost or damaged myocardium in a mammal, or to prophylactically treat a mammal at risk of such loss or damage. The present invention is novel in that myocardial tissue is believed to lack a sufficient number of myogenic precursor cells for adequate regeneration or repair of lost or damaged tissue and, therefore, the ability of the

morphogens to promote the migration, proliferation and/or differentiation of myogenic precursor cells (e.g., skeletal muscle satellite cells) into functional myocardium is unexpected.

In preferred embodiments, the morphogen is a dimeric protein comprising a pair of folded polypeptides, each having an amino acid sequence that shares a defined relationship with an amino acid sequence of a reference morphogen. Preferred morphogen polypeptides share a defined relationship with a sequence present in morphogenically active human OP-1 (SEQ ID NO: 4). However, any one or more of the naturally occurring or biosynthetic sequences disclosed herein similarly could be used as a reference sequence. Preferred morphogen polypeptides share a defined relationship with at least the C-terminal six cysteine domain of human OP-1 (residues 43-139 of SEQ ID NO: 4). Preferably, morphogen polypeptides share a defined relationship with at least the C-terminal seven cysteine domain of human OP-1 (residues 38-139 of SEQ ID NO: 4). That is, preferred morphogen polypeptides in a dimeric protein with morphogenic activity each comprise a sequence that corresponds to a reference sequence or is functionally equivalent thereto. Examples of preferred morphogens include mammalian, and particularly human, OP-1, CBMP-2A (BMP-2) and CBMP-2B (BMP-4).

Functionally equivalent sequences include functionally equivalent arrangements of cysteine residues disposed within the reference sequence, including amino acid insertions or deletions which alter the linear arrangement of these cysteines, but do not materially impair their relationship in the folded structure of the dimeric morphogen protein, including their ability to form such intra- or inter-chain disulfide bonds as may be necessary for morphogenic activity. Functionally equivalent sequences further include those wherein one or more amino acid residues differs from the corresponding residue of a reference morphogen sequence, e.g., the C-terminal seven cysteine domain (also referred to herein as the conserved seven cysteine skeleton) of human OP-1, provided that this difference does not destroy morphogenic activity. Accordingly, conservative substitutions of corresponding amino acids in the reference sequence are preferred. Amino acid residues that are "conservative substitutions" for corresponding residues in a reference sequence are those that are physically or functionally similar to the corresponding reference residues, e.g., that have similar size, shape, electric charge, chemical properties including the ability to form covalent or hydrogen bonds, or the like. Particularly preferred conservative substitutions are those fulfilling the criteria defined for an "accepted point mutation"

in Dayhoff, et al. (1978) Atlas of Protein Sequence and Structure, 5: Suppl. 3, ch. 22 (pp. 354-352), Natl. Biomed. Res. Found., Washington, D.C. 20007, the teachings of which are incorporated by reference herein.

In certain embodiments, a polypeptide suspected of being functionally equivalent to a reference morphogen polypeptide is aligned therewith using the method of Needleman, et al. (1970) J. Mol. Biol. 48:443-453, implemented conveniently by computer programs such as the Align program (DNAstar, Inc.). As noted above, internal gaps and amino acid insertions in the candidate sequence are ignored for purposes of calculating the defined relationship, conventionally expressed as a level of amino acid sequence homology or identity, between the candidate and reference sequences. "Amino acid sequence homology" is understood herein to include both amino acid sequence identity and similarity. Homologous sequences share identical and/or similar amino acid residues, where similar residues are conservative substitutions for, or "allowed point mutations" of, corresponding amino acid residues in an aligned reference sequence. Thus, a candidate polypeptide sequence that shares 70% amino acid homology with a reference sequence is one in which any 70% of the aligned residues are either identical to, or are conservative substitutions of, the corresponding residues in a reference sequence.

The present invention alternatively can be practiced with methods and compositions comprising a morphogen inducer in lieu of a morphogen. A "morphogen inducer" is a compound that stimulates the production (i.e., transcription, translation, and/or secretion) of morphogen by a cell competent to produce and/or secrete a morphogen encoded within the genome of the cell. Endogenous or administered morphogens can act as endocrine, paracrine or autocrine factors. Therefore, an inducer of a morphogen may stimulate endogenous morphogen synthesis by the cells in which the morphogenetic responses are induced, by neighboring cells in vivo or in vitro (e.g., in tissue culture) or by cells of a distant tissue in vivo (in which case the secreted morphogen is transported to the site of morphogenesis, e.g., by the individual's bloodstream). In preferred embodiments, the inducer stimulates expression and/or secretion of a morphogen so as to increase amounts thereof available to mammalian myogenic precursor cells in vivo or in vitro. Thus, to promote the migration, proliferation and/or differentiation of myogenic precursor cells in vivo, an inducer of a morphogen may be administered to induce production of morphogen by the myogenic precursor cells themselves, or by other cells co-cultured with the myogenic precursor

cells. Similarly, to promote the proliferation and/or differentiation of myogenic precursor cells in vivo, an inducer of a morphogen may administered locally or systemically to induce morphogen production by the myogenic precursor cells themselves, or by neighboring or distant cells in a mammal's body.

In still other embodiments, an agent which acts as an agonist of a morphogen receptor may be administered instead of the morphogen itself. An "agonist" of a receptor is a compound which binds to the receptor, and for which the result of such binding is similar to the result of binding the natural, endogenous ligand of the receptor. That is, the compound must, upon interaction with the receptor, produce the same or substantially similar transmembrane and/or intracellular effects as the endogenous ligand. Thus, an agonist of a morphogen receptor binds to the receptor and such binding has the same or a functionally similar result as morphogen binding (e.g., induction of morphogenesis). The activity or potency of an agonist can be less than that of the natural ligand, in which case the agonist is said to be a "partial agonist," or it can be equal to or greater than that of the natural ligand, in which case it is said to be a "full agonist." Thus, for example, a small peptide or other molecule which can mimic the activity of a morphogen in binding to and activating the morphogen's receptor may be employed as an equivalent of the morphogen. Preferably the agonist is a full agonist, but partial morphogen receptor agonists may also be advantageously employed. Methods of identifying such agonists are known in the art and include assays for compounds which induce morphogen-mediated responses (e.g., induction of differentiation of metanephric mesenchyme, induction of endochondral bone formation, and the like). Such an agonist may also be referred to as a morphogen "mimic," "mimetic," or "analog."

Alternatively, a small molecule morphogenic activator, as described herein, may be administered instead of the morphogen itself to promote the migration, proliferation, and/or differentiation of myogenic precursor cells by increasing the level of expression of proteins associated with myocardial phenotype. Exemplary methods comprise introducing a small molecule morphogenic activator that regulates some portion or portions of a morphogen-induced regulatory pathway, resulting in an effective increase in expression or activity of myocardium-specific protein. This may result either from stimulating an increase in the endogenous expression of such protein or from a decrease in the inhibition of normal expression of such protein. For example, a small molecule morphogenic activator may act at the type I or type II morphogen

receptor; or at the serine/threonine kinase, or other kinase domains of those receptors. Another target of pathway activation is the Smad proteins, including the monomeric, dimeric (including heteromeric and homomeric complexes) or trimeric forms (including heteromeric and homomeric complexes). Alternately, a small molecule morphogenic activator may lead to activation of a transcription factor (for example, the X-protein shown in Figure 2) that causes phenotype-specific gene expression (i.e., expression of protein characteristic of myocardium).

Preferably, the morphogens, morphogen inducers, agonists of morphogen receptors, or small molecule morphogenic activators are directly contacted with the myogenic precursor cells in solution either in vitro prior to implantation, in vivo at the time of implantation, or in vivo subsequent to implantation. Alternatively, however, the morphogens, morphogen inducers, agonists of morphogen receptors may be administered by any route which is compatible with the selected agent, and may be formulated with any pharmaceutically acceptable carrier appropriate to the route of administration. Preferred systemic routes of administration are parenteral and, in particular, intravenous and intraperitoneal.

In additional embodiments, the present invention provides pharmaceutical compositions comprising a morphogen, or morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator in combination with one or more of a "muscle extract," conditioned medium from differentiated myotubes grown in culture, bFGF, IGF, PDGF, LIF, ACTH, MSH, or G-CSF. These compositions are useful in promoting the proliferation and/or differentiation of myogenic precursor cells.

Brief Description of the Figures

Figure 1. Panels 1-1 through 1-12 of this figure are a tabular alignment of the amino acid sequences of various naturally occurring morphogens with a preferred reference sequence of human OP1, residues 38-139 of SEQ ID NO: 4. Morphogen polypeptides shown in this figure also are identified in the Sequence Listing.

Figure 2. Figure 2 is a schematic representation of a morphogen-activated regulatory pathway for expression of a phenotype-specific gene.

Detailed Description of the Invention

I. Definitions

In order to more clearly and concisely point out the subject matter of the claimed invention, the following definitions are provided for specific terms used in the following written description and appended claims.

Subjects at risk of, or afflicted with, loss of or damage to myocardium. As used herein, a subject (preferably a mammal, e.g., a human) is said to be at risk of, or afflicted with, loss of or damage to myocardium, if the subject has suffered a loss of functional myocardial tissue which is clinically detectable in terms of reduced or altered cardiac function, or if the subject may reasonably be expected to suffer such a loss. Subjects at risk of, or afflicted with, loss of or damage to myocardium include, but are not limited to, subjects which have already suffered a myocardial infarction, which have suffered a physical trauma to the heart (e.g., in an automobile accident) which has reduced cardiac function, or which have already been diagnosed with congestive heart failure; as well as subjects which can reasonably be expected to suffer a myocardial infarction or congestive heart failure. Whether a particular subject is at risk is a determination which may routinely be made by one of ordinary skill in the relevant medical or veterinary art.

Myogenic precursor cells. As used herein, the term "myogenic precursor cells" refers to cells capable of myogenesis, or the process of proliferation and differentiation into new and functional muscle when present in a morphogenically permissive environment. Myogenic precursor cells are variously referred to in the literature as "myoblasts," "muscle stem cells" or "satellite cells."

Morphogenically permissive environment. As used herein, a "morphogenically permissive environment" is an environment which allows or promotes the differentiation of cells into a specific cell type or types. A "morphogenically permissive environment" is, therefore, sufficiently free of inhibitors of cell differentiation to allow or promote cell differentiation. In addition, a morphogenically permissive environment is one which provides signals (e.g., through cell-cell contact, cell-extracellular matrix contact, or diffusible factors) which allow or promote a pluripotent cell to follow a particular morphogenic pathway. In particular, with respect to myocardial differentiation, a morphogenically permissive environment includes an environment of intact or damaged myocardial tissue which provides signals to myogenic precursor cells which allow or promote the differentiation of those cells into new and functional myocardium. It is

known, for example, that myogenic precursor cells differentiate into myocytes at least partly in response to contact with the plasmalemma of a myofiber. The presence of myofiber plasmalemmas, therefore, may be one element of a morphogenically permissive environment for myogenesis. Similarly, electrical or biochemical stimuli from nerves, as well as a variety of growth factors (see below), appear to be elements of a morphogenically permissive environment for myogenesis. Thus, a morphogenically permissive environment may include one or more of these elements.

II. Description of the Preferred Embodiments

A. General

The present invention depends, in part, upon the surprising discovery that morphogenically-treated mammalian myogenic precursors cells, when implanted in vivo at a site of lost or damaged mammalian myocardium, undergo a process of proliferation and/or differentiation to produce new and functional mammalian myocardium, thereby restoring or regenerating the lost or damaged tissue in whole or in part. This result is particularly unexpected in light of the fact that mammalian myocardial tissue is believed to lack a sufficient number of myogenic precursor cells for adequate regeneration or repair of lost or damaged tissue and, therefore, mammalian myocardium previously has been believed to be a poor responder for functional restoration or regeneration after tissue loss or damage. In addition, the present invention depends, in part, upon the surprising discovery that non-myocardial cells, such as those obtained from mammalian skeletal muscle or embryonic myogenic precursor cells, may be induced to proliferate and differentiate into myocardium in a morphogenically permissive environment. It is further surprising that the morphogens, morphogen inducers, agonists of morphogen receptors, and small molecule morphogenic activators, as described herein, may promote such restoration or regeneration despite the fact that they have no known role in myocardial tissue restoration or regeneration in the adult mammal.

Without being bound to any particular theory of the invention, it is believed that the morphogens, morphogen inducers, agonists of morphogen receptors, or small molecule morphogenic activators may promote the proliferation of myogenic precursor cells and render them more susceptible to differentiation into new and functional myocardium when implanted in a morphogenically permissive environment. Thus, it is believed that the morphogens, morphogen

inducers, agonists of morphogen receptors, or small molecule morphogenic activators may increase the pluripotentiality of these myogenic precursor cells, such that they may "switch fates" and, rather than differentiating only into smooth or skeletal muscle, they may proliferate and then differentiate into new and functional myocardium.

B. Isolating and Culturing Mammalian Myogenic Precursor Cells

Methods of isolating and culturing mammalian myogenic precursor cells are well-established in the art. For example, myogenic precursor cells may be obtained, as further described in the examples below, by dissociation of skeletal muscle and subsequent culturing of the satellite cells. Alternatively, myogenic precursor cells may be obtained from embryonic tissues, where they arise as fetal myoblasts from the myotomes of the somites, after induction of the mesoderm. Myogenic precursor cells may also be obtained from cell lines, such as a pluripotent mesodermal mesenchyme cell line or a partially dedifferentiated laboratory cell line, which may be induced to differentiate into myoblasts after implantation into a morphogenically permissive environment. See, generally, Hathaway, et al. (1991) J. Cell. Physiol. 146:435-441; Mezzogiorno et al. (1993) Mech. Ageing & Develop. 70:35-44; Alameddine and Fardeau (1989); Chiu et al. (1995) Ann. Thorac. Surg. 60:12-18.

1. Isolating Myogenic Precursor Cells from Skeletal Muscle

In preferred embodiments, the myogenic precursor cells are obtained from skeletal muscle. The skeletal muscle donor is preferably the subject for myocardial treatment or an identical twin in order to avoid problems of histocompatibility and possible tissue rejection. Alternatively, other family members or histocompatible donors, including transgenic mammals raised for organ transplantation purposes (e.g., lacking MHC markers or expressing humanized MHC proteins), may be employed as donors of the skeletal muscle tissue. Depending upon the degree of histocompatibility, standard methods of immunosuppression may be needed in conjunction with the present invention to prevent rejection of the implanted cells.

Briefly, a sample of skeletal muscle is excised from one or more skeletal muscles of a subject under local or general anesthesia. Any excessive connective tissue and fasciae are dissected away, the muscle is rinsed in sterile solution, and the muscle is dissociated by, for example, mincing with scissors or passage through a meat grinder until substantially homogeneous. The amount of muscle excised will depend, of course, upon the quantity of

myogenic precursor cells required by the treatment, as well as the degree of myogenic precursor cell proliferation which is to be promoted in vitro. Typically, however, amounts of 1-100 grams, more preferably 10-50 grams, of skeletal muscle tissue are removed. Such quantities may be excised conveniently from one or more of the larger, relatively superficial muscles of the limbs (e.g., biceps brachii, triceps brachii, brachialis, brachioradialis, rectus femoris, biceps femoris, semitendinosus, gracilis, vastus lateralis, gastrocnemius, tibialis anterior), chest and shoulders (e.g., pectoralis, deltoid), pelvis and hips (e.g., gluteus medius, gluteus maximus), back (e.g., trapezius, latissimus dorsi) or abdomen (e.g., obliquus abdominis externus, rectus abdominis), but may be obtained from any available skeletal muscle.

Preferably, the dissociated muscle then is incubated with a proteolytic enzyme (e.g., pronase (Sigma, St. Louis, MO), collagenase (Sigma, St. Louis, MO), hyaluronidase (Sigma, St. Louis, MO), or trypsin (Difco Laboratories, Inc., Detroit, MI) at 37°C for 15 min to 1 hr to remove remaining connective tissue. The mass of digested muscle tissue optionally may be further dissociated by, for example, repeated pipetting or mixing. In addition, the digested mass optionally may be washed, pelleted and resuspended to remove digested connective tissue and enzyme, and any remaining debris may be removed by filtration. The cells are then suspended in a sterile buffer (e.g., phosphate buffered saline solution) and centrifuged at approximately 500-550 g for approximately 10 minutes to sediment the larger, multinucleated skeletal muscle fibers and myocytes, while leaving the satellite cells in the supernatant. Either before or after centrifugation, serum, such as fetal bovine serum (FBS, GIBCO BRL, Grand Island, NY), may be added to the mixture to halt the enzymatic cleavage process and antibiotics may be added to prevent microbial growth. If desired, satellite cells may be separated from fibroblasts and other remaining cells using a density centrifugation method (see, e.g., Yablonka-Reuveni and Nameroff (1987) Histochemistry 87:27-38).

2. Isolating Myogenic Precursor Cells from Embryos

Myogenic precursors cells may be isolated from mammalian embryonic or fetal (together "embryonic") tissues at various stages of development after induction of the mesoderm. Thus, for example, myogenic precursor cells may be obtained from the embryonic mesoderm prior to its further differentiation into dorsal, intermediate, and lateral mesodermal mesenchyme. After this stage of differentiation, any mesodermal cells may be employed but, preferably, cells are employed

which arise along the routes of differentiation toward skeletal or cardiac muscle. For example, the dorsal mesodermal mesenchyme differentiates to form the myotomes which, in turn, differentiate to form both the skeletal muscles of the trunk and the limb buds. The mesodermal mesenchyme of the limb buds further differentiates to form the skeletal muscles of the appendages (as well as the appendicular skeleton. Similarly, the lateral mesodermal mesenchyme differentiates, in part, to form the splanchic mesoderm which, in turn, differentiates to form the myocardium and smooth muscles of the viscera (as well as the gonads, circulatory system and other primary elements of the viscera). One of ordinary skill in the art may, therefore, readily choose appropriate embryonic cells for use in the present invention (see, e.g., Soonpaa et al. (1994) Science 264:98-101; also see, generally, B.M. Carlson, ed. (1981) Patten's Foundations of Embryology, 4th Edition, McGraw-Hill, New York). Once excised, the embryonic tissue may be treated essentially as described above with respect to skeletal muscle to isolate the myogenic precursor cells.

As with cells obtained from the skeletal muscle of an adult mammal, histocompatibility problems may arise upon implantation of embryonic myogenic precursor cells. Therefore, depending upon the degree of histocompatibility, standard methods of immunosuppression may be needed in conjunction with the present invention to prevent rejection of the implanted cells.

3. Isolating Myogenic Precursor Cells from Established Cell Lines

Established cell lines, including myogenic precursor cell lines, myoblast cell lines, or mesenchymal cell lines, may also be employed in the present invention without the need for isolation of the myogenic precursor cells from adult or embryonic tissue. For example, the established murine myoblast cell line C_2C_{12} (ATCC CRL 1772) has been implanted into mouse hearts and shown to differentiate into functional myocardium and fuse with native myocardium (Koh et al. (1993) J. Clin. Invest. 92:1548-54). Alternatively, pluripotent mesodermal stem cell lines, including primary dermal fibroblast lines, smooth muscle cell lines, or chondroblast lineages may be caused to differentiate into muscle cells (see, e.g., Choi et al. (1990) Proc. Nat. Acad. Sci. (USA) 87:7988-7992). Finally, it should be noted that a variety of established mammalian myogenic cell lines are commercially available for use in accordance with the present invention including, for example, the human cell line HISM (ATCC CRL 1692), the murine cell lines C2C12 (ATCC CRL 1772), NOR-10 (ATCC CRL 197), and G-8 (ATCC CRL 1456), and the rat

cell lines A7r5 (ATCC CRL 1444), A10 (ATCC CRL 1476), H9c2 (2-1) (ATCC CRL 1446), L6 (ATCC CRL 1458) and L8 (ATCC CRL 1769). Following essentially the same protocols as described in the original reports of these cell lines (see the ATCC's Catalogue of Cell Lines & Hybridomas, for citations) one of ordinary skill in the art can readily produce comparable cell lines from any mammalian species.

4. Culturing Myogenic Precursor Cells

Myogenic precursor cells may be cultured on solid or in liquid media. Thus, for example, the myogenic precursor cells may be suspended in a flask of liquid medium while maintaining mild or periodic agitation. Alternatively, the cells may be plated on a solid substrate and fed with a liquid medium. Appropriate liquid media are well known in the art and include, but are not limited to, McCoy's, M199, Minimal Essential Medium (MEM), Dulbecco's Modified Eagle Medium (commercially available from, for example, GIBCO BRL, Grand Island, NY, or Sigma Chemical Company, St. Louis, MO), and the like. These media may, of course, be supplemented with additional buffers or nutrient solutions (e.g., 10% fetal bovine serum, 3% horse serum), or with antimycotics and/or antibiotics (e.g., 50-5,000 IU/ml penicillin, 50-5,000 µg/ml streptomycin, 5-50 µg/ml gentamicin). Preferably, the liquid media is replaced every 24-48 hrs and the cultures are maintained at a relatively constant temperature of about 37°C under a normal or 5% CO_2-enriched humid atmosphere. For culturing on solid substrates, cells are preferably plated at a density of approximately 10^4-10^6 cells per 60 mm plate. To promote cell adherence to solid substrates, the plates may optionally be coated with, for example, basement membrane matrigel or laminin (Sigma Chemical Company, St. Louis, MO) although, as described below, adherence and/or confluence may inhibit proliferation.

In order to allow or promote proliferation of the myogenic precursor cells in vitro while inhibiting premature differentiation, a number of steps may be taken. For example, myogenic precursor cell proliferation has been shown to be inhibited by TGF-β (Allen and Boxhorn (1989) J. Cell Physiol. 138:311-315) and contact with myofiber plasmalemmas, (Bischoff (1989)); and has been shown to be promoted by a saline "muscle extract" (Bischoff (1986) Dev. Biol. 115:140), conditioned medium from differentiated myotubes grown in culture (Mezzogiorno et al. (1993) Mech. Ageing & Develop. 70:35-44), basic fibroblast growth factor (bFGF) (Clegg et al. (1987) J. Cell. Biol. 105:949-56), insulin-like growth factors (IGF) (Ewton and Florini (1977)

Endocrinology 106:577-587; Tollfsen et al. (1989) Proc. Nat. Acad. Sci. (USA) 1543-1547), platelet-derived growth factor (PDGF) (Yablonka-Reuveni et al. (1990) J. Cell Biol. 111:1623-1629), leukemia-inhibiting factor (LIF) (Austin and Burgess (1991) J. Neuro. Sci. 101:193-197), adrenocorticotrophic hormone (ACTH) (Cossu et al. (1989) Develop. Biol. 131:331-336; De Angelis et al. (1992) Dev. Biol. 151:446-458), melanocyte-stimulating factor (MSH) (Cossu et al. (1989) Develop. Biol. 131:331-336) and granulocyte colony stimulating factor (G-CSF) (Austin and Burgess (1991) J. Neuro. Sci. 101:193-197). Thus, in order to promote proliferation of the myogenic precursors cells in vitro prior to implantation and/or in vivo after implantation, the cells may be grown in the presence of one or more of these factors, or other known mitogens. In addition, as is generally known in the art, proliferation of such cells may be promoted by repeated passaging (e.g., treatment with dilute trypsin to remove adhered cells from the culture plate and replating at a lower density every 2-3 days), growth in liquid culture, growth in the absence of enhancers of cell adhesion, growth in the presence of inhibitors of cell adhesion, and/or growth at densities below confluence.

There is no absolute requirement that the myogenic precursor cells of the present invention be cultured in vitro prior to implantation. Indeed, if a therapeutically effective number of myogenic precursor cells can conveniently and economically be obtained without culturing, this step may be omitted. On the other hand, when such cells are in scarce supply (e.g., from fetal tissues) or can be obtained only through invasive measures (e.g., excision of substantial portions of muscle from a donor or donor/subject), it is preferred that smaller numbers of cells be obtained initially, and then proliferated in vitro. Doubling times will vary depending upon the source of cells, media, and the presence or absence of other growth factors, but doubling times on the order of every 12 hrs have been reported in the literature for muscle satellite cells grown in the presence of muscle abstract (Bischoff, (1989)). Therefore, it is contemplated that culturing times of several days to a week may be employed in the present methods to expand the myogenic precursor cell population prior to implantation.

Myogenic precursor cells may be harvested by brief trypsin treatment to remove any cells adhered to the culture plate or vessel, and centrifugation (e.g., 10-15 min at 500-1000 g). The cells may then be resuspended in a physiologically acceptable buffer solution (e.g., PBS, Ringer's saline) at an appropriate density (e.g., 10^3-10^7 cells/ml).

Finally, it should be noted that morphogens, morphogen inducers, agonists of morphogen receptors, and small molecule morphogenic activators may be used to treat the myogenic precursor cells during culturing (if any) to aid in proliferation and/or subsequent differentiation. Alternatively, the myogenic precursor cells may be treated with a morphogen, morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator either simultaneously with, or subsequent to, implantation. In the case of morphogen inducers, the myogenic precursor cells may be co-cultured with auxiliary cells which respond to these morphogen inducers by producing morphogen. The myogenic precursor cells then may be implanted along with these auxiliary cells, or may be isolated from the co-culture by standard cell separation techniques, which are known in the art, but which will vary with the type of auxiliary cells employed (e.g., density centrifugation separation, cell type specific cytotoxins).

C. Implantation of Myogenic Precursor Cells at a Myocardial Site

Myogenic precursor cells may be implanted at a site of loss of or damage to mammalian myocardium by any of a variety of surgical techniques known in the art. These techniques range from the minimally invasive (e.g., injection by needle through the thoracic wall) to substantially invasive (e.g., thoracotomy and incision of the myocardium, followed by implantation, suturing of the implant site and closing of the chest). The technique employed in any given instance will depend upon such factors as the size of the myocardial site to be treated, the accessibility of the site, and the age and stamina of the subject.

Generally, the myogenic precursor cells are implanted in a physiologically acceptable buffer solution. To minimize the volume of solution administered to the treatment site, the cells may be at a relatively high titer within this solution (e.g., 10^5-10^7 cells/ml). The solution may contain growth factors, as described above, to promote further proliferation of the myogenic precursor cells within the implant site, or may be free of such factors so as to promote differentiation into new and functional myocardium in the morphogenically permissive environment of the myocardial implant site. In addition, as noted above, the myogenic precursor cells may be implanted either simultaneously with a morphogen, morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator, or the morphogenic treatment may be subsequent to implantation.

Thus, for example, a solution of myogenic precursor cells and a morphogen, morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator, may be implanted at a site of myocardial infarction in essentially the following manner. For example, to treat a myocardial infarct to the anterior wall of the left ventricle, a left thoracotomy is performed on a subject under general anesthesia in an intercostal space (e.g., the sixth intercostal space) and the site of the infarct is determined by observation. At the discretion of the surgeon, the heart may or may not be stopped and systemic blood flow shunted to a heart-lung machine. Myogenic precursor cells then may be directly injected into one or more sites within the infarct using an intravenous catheter (e.g., a 16-gauge Teflon catheter from Criticon, Tampa, FL). The initial injection(s) may include a morphogen, morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator, or these may be included in one or more additional injections to the infarct site. Alternatively, a number of non-transmural incisions may be made at the site of the infarct to create "channels" parallel to the direction of the myocardial fibers. The suspension of myogenic precursor cells (with or without morphogen, morphogen inducer or morphogen receptor agonist) then may be introduced within these channels and the channels closed by suturing. Finally, the pericardium is sutured and chest wall are closed by standard surgical techniques (after restarting and returning systemic circulation to the heart from a heart-lung machine, if employed).

The treatment of chronically deteriorating mammalian myocardium (e.g., due to congestive heart failure or chronic myopathy), may be performed similarly except that the implantation sites are chosen to correspond to areas of generalized myocardial deterioration and, therefore, may be more diffuse.

The number of myogenic precursor cells implanted will vary according to the amount of myocardial tissue to be restored or regenerated. The volume of cells to be restored or regenerated may be ascertained by standard techniques of cardiac imaging. Generally, it is expected that on the order of approximately 10^4-10^5 myogenic precursor cells will be required to restore or regenerate 1 mg of myocardial tissue (see, e.g., Alameddine and Fardeau (1989)).

D. Morphogens, Inducers, Agonists, and Small Molecule Morphogenic Activators

Morphogens useful in the present invention include eukaryotic proteins originally identified as osteogenic proteins (see U.S. Patent 5,011,691, incorporated herein by reference),

such as the OP1, OP2, OP3, CBMP2A (BMP-2), CBMP-2B (BMP-4) and BMP3 proteins (SEQ ID NOs: 4-9, 15-22, 25-27), as well as amino acid sequence-related proteins such as DPP (SEQ ID NO: 10, from Drosophila), Vgl (SEQ ID NO: 11, from Xenopus), Vgr1 (SEQ ID NO: 12, from mouse), GDF1 (SEQ ID NOs: 13, 30 and 31, from humans, see Lee (1991), PNAS 88:4250-4254), 60A (SEQ ID NOs: 23 and 24, from Drosophila, see Wharton et al. (1991) PNAS 88:9214-9218), dorsalin-1 (from chick, see Basler et al. (1993) Cell 73:687-702 and GenBank accession number L12032) and GDF5 (from mouse, see Storm et al. (1994) Nature 368:639-643). Additional useful morphogens include biosynthetic morphogen constructs disclosed in U.S. Pat. No. 5,011,691, e.g., COP1, 3-5, 7 and 16, as well as others known in the art including dor3, NODAL, UNIVIN, BMP9, BMP10, GDF3, GDF6, GDF7, CDMP2, and SCREW. See also U.S. Pat. No. 4,968,590, incorporated herein by reference.

Naturally occurring proteins identified and/or appreciated herein to be morphogens form a distinct subgroup within the loose evolutionary grouping of sequence-related proteins known as the TGFβ superfamily or supergene family. The naturally occurring morphogens share substantial amino acid sequence homology in their C-terminal regions (domains). Typically, the above-mentioned naturally occurring morphogens are translated as a precursor, having an N-terminal signal peptide sequence, typically less than about 30 residues, followed by a "pro" domain that is cleaved to yield the mature C-terminal domain. The signal peptide is cleaved rapidly upon translation, at a cleavage site that can be predicted in a given sequence using the method of Von Heijne (1986) Nucleic Acids Research 14:4683-4691. The pro domain typically is about three times larger than the fully processed mature C-terminal domain. Herein, the "pro" form of a morphogen refers to a morphogen comprising a folded pair of polypeptides each comprising the pro and mature domains of a morphogen polypeptide. Typically, the pro form of a morphogen is more soluble than the mature form under physiological conditions. The pro form appears to be the primary form secreted from cultured mammalian host cells.

Table 1, below, summarizes various naturally occurring morphogens identified to date, including their nomenclature as used herein, their Sequence Listing references, and publication sources for the amino acid sequences for the full length proteins not included in the Sequence Listing. Each of the generic terms set forth in Table 1 is intended and should be understood to embrace morphogenically active proteins expressed from nucleic acids encoding the identified

sequence mentioned below and set forth in the Sequence Listing, or a morphogenically active fragment or precursor thereof, including functional equivalents such as naturally occurring and biosynthetic variants thereof. Naturally occurring variants include allelic variant forms isolated from other individuals of a single biological species, and phylogenetic counterpart (species) variant forms (homologues) isolated from phylogenetically distinct biological species. The disclosures of publications mentioned below are incorporated herein by reference.

TABLE 1

"OP1"	Refers generically to morphogenically active proteins expressed from nucleic acids encoding OP1 proteins, including at least the human OP1 protein disclosed in SEQ ID NO: 4 ("hOP1"), and the mouse OP1 protein disclosed in SEQ ID NO: 5 ("mOP1"). In each of human and mouse OP1 proteins, the conserved seven cysteine skeleton is defined by residues 38 to 139. cDNA sequences and amino acid sequences encoded therein and corresponding to the full length proteins are provided in SEQ ID NOs: 15 and 16 (hOP1) and SEQ ID NOs: 17 and 18 (mOP-1.) The mature proteins are defined by residues 293-431 (hOP1) and 292-430 (mOP1). The "pro" regions of the proteins, cleaved to yield the mature, morphogenically active proteins are defined essentially by residues 30-292 (hOP1) and residues 30-291 (mOP1).
"OP2"	Refers generically to morphogenically active proteins expressed from nucleic acids encoding the OP2 proteins, including at least the human OP2 protein disclosed in SEQ ID NO: 6 ("hOP2"), and the mouse OP2 protein disclosed in SEQ ID NO: 7 ("mOP2"). In each of human and mouse OP2 proteins, the conserved seven cysteine skeleton is defined by residues 38 to 139 of SEQ ID NOs: 6 and 7. cDNA sequences and amino acid sequences encoded therein and corresponding to the full length proteins are provided in SEQ ID NOs: 19 and 20 (hOP2) and SEQ ID NOs: 21 and 22 (mOP2.) The mature proteins are defined essentially by residues 264-402 (hOP2) and 261-399 (mOP2). The "pro" regions of the proteins, cleaved to yield the mature, morphogenically active proteins are defined essentially by residues 18-263 (hOP2) and residues 18-260 (mOP1).

"OP3" Refers generically to morphogenically active proteins expressed from nucleic acids encoding OP3 proteins, including at least the mouse OP3 protein disclosed in SEQ ID NO: 26 ("mOP3"). The conserved seven cysteine domain is defined by residues 298 to 399 of SEQ ID NO: 26, which shares greater than 79% amino acid identity with the corresponding mOP2 and hOP2 sequences, and greater than 66% identity with the corresponding OP1 sequences. A cDNA sequence encoding the above-mentioned amino acid sequence is provided in SEQ ID NO: 25. OP3 is unique among the morphogens identified to date in that the residue at position 9 in the conserved seven cysteine domain (e.g., residue 315 of SEQ ID NO: 26) is a serine, whereas other morphogens typically have a tryptophan at this location.

"CBMP2" Refers generically to morphogenically active proteins expressed from nucleic acids encoding the CBMP2 proteins, including at least the human CBMP2A protein disclosed in SEQ ID NO: 8 (hCBMP2A) and the human CBMP2B protein disclosed in SEQ ID NO: 9 (hCBMP2B). The amino acid sequence for the full length proteins, referred to in the literature as BMP2A and BMP2B, or BMP2 and BMP4, appear in Wozney, et al. (1988) Science 242:1528-1534. The pro domain for BMP2 (BMP2A) likely includes residues 25-248 of the published sequence; the mature protein, residues 249-396. The pro domain for BMP4 (BMP2B) likely includes residues 25-256 of the published sequence; the mature protein, residues 257-408.

"DPP" Refers generically to proteins encoded by the Drosophila DPP gene and defining at least the conserved seven cysteine skeleton of SEQ ID NO: 10. The amino acid sequence for the full length protein appears in Padgett, et al. (1987) Nature 325:81-84. The pro domain likely extends from the signal peptide cleavage site to residue 456 of the published sequence; the mature protein likely is defined by residues 457-588.

"Vgl" Refers generically to proteins encoded by the Xenopus Vgl gene and defining at least the conserved seven cysteine skeleton of SEQ ID NO: 11. The amino acid sequence for the full length protein appears in Weeks (1987) Cell 51:861-867. The

prodomain likely extends from the signal peptide cleavage site to residue 246 of the published sequence; the mature protein likely is defined by residues 247-360.

"Vgr1" Refers generically to proteins encoded by the murine Vgr1 gene and defining at least the conserved seven cysteine skeleton of SEQ ID NO: 12. The amino acid sequence for the full length protein appears in Lyons, et al. (1989) PNAS 86:4554-4558. The prodomain likely extends from the signal peptide cleavage site to residue 299 of the published sequence; the mature protein likely is defined by residues 300-438.

"GDF1" Refers generically to proteins encoded by the human GDF1 gene and defining at least the conserved seven cysteine skeleton of SEQ ID NO: 13. The cDNA and encoded amino sequence for the full length protein are provided in SEQ ID NOs: 30 and 31. The prodomain likely extends from the signal peptide cleavage site to residue 214; the mature protein likely is defined by residues 215-372.

"60A" Refers generically to morphogenically active proteins expressed from nucleic acid encoding 60A proteins or morphogenically active fragments thereof, including at least the Drosophila 60A protein disclosed in SEQ ID NO: 24. A Drosophila 60A cDNA is disclosed in SEQ ID NO: 23. The prodomain likely extends from the signal peptide cleavage site to residue 324; the mature protein likely is defined by residues 325-455. The active fragment of 60A protein likely is defined by the conserved seven cysteine skeleton of residues 354 to 455 of SEQ ID NO: 24. The 60A protein is considered likely herein to be a phylogenetic counterpart variant of the human and mouse OP1 genes; Sampath, et al. (1993) PNAS 90:6004-6008.

"BMP3" Refers generically to proteins encoded by the human BMP3 gene and defining at least the conserved seven cysteine skeleton of SEQ ID NO: 27. The amino acid sequence for the full length protein appears in Wozney, et al. (1988) Science 242:1528-1534. The pro domain likely extends from the signal peptide cleavage

site to residue 290 of the published sequence; the mature protein likely is defined by residues 291-472.

"BMP5" Refers generically to proteins encoded by the human BMP5 gene and defining at least the conserved seven cysteine skeleton of SEQ ID NO: 28. The amino acid sequence for the full length protein appears in Celeste, et al. (1991) PNAS 87:9843-9847. The pro domain likely extends from the signal peptide cleavage site to residue 316 of the published sequence; the mature protein likely is defined by residues 317-454.

"BMP6" Refers generically to proteins encoded by the human BMP6 gene and defining at least the conserved seven cysteine skeleton of SEQ ID NO: 29. The amino acid sequence for the full length protein appears in Celeste, et al. (1990) PNAS 87:9843-5847. The pro domain likely extends from the signal peptide cleavage site to residue 374 of the published sequence; the mature protein likely is defined by residues 375-513.

As shown in Figure 1, the OP2 and OP3 proteins have an additional cysteine residue in the conserved C-terminal region (e.g., see residue 41 of SEQ ID NOs: 6 and 7), in addition to the conserved cysteine skeleton or domain in common with the other known proteins in this family. The GDF1 protein has a four amino acid insert within the conserved skeleton (residues 44-47 of SEQ ID NO: 13) but this insert likely does not interfere with the relationship of the cysteines in the folded structure. Further, the CBMP2 proteins are missing one amino acid residue within the cysteine skeleton. Thus, these morphogen polypeptides illustrate the principles of alignment used herein with respect to the preferred reference morphogen sequence of human OP1, residues 38-139 of SEQ ID NO: 4.

In certain preferred embodiments, morphogens useful herein include those in which the amino acid sequences of morphogen polypeptides comprise a sequence sharing at least 70% amino acid sequence homology or "similarity", and preferably 80% homology or similarity with a reference morphogen sequence selected from the foregoing sequences or naturally occurring morphogens. Preferably, the reference morphogen is human OP1, and the reference sequence thereof is the C-terminal seven cysteine domain present in morphogenically active forms of human

OP1, residues 38-139 of SEQ ID NO: 4. Morphogens useful herein accordingly include alleles, phylogenetic counterparts and other variants of the preferred reference sequence, whether naturally-occurring or biosynthetically produced (e.g., including "muteins" or "mutant proteins"), as well as novel members of the morphogenic family of proteins including the morphogens set forth and identified above, e.g., in connection with Table 1. Certain particularly preferred morphogen polypeptides share at least 60% amino acid identity with the preferred reference sequence of human OP1, still more preferably at least 65% amino acid identity therewith.

In other preferred embodiments, the family of morphogen polypeptides useful in the present invention, and members thereof, are defined by a generic amino acid sequence. For example, Generic Sequence 7 (SEQ ID NO: 1) and Generic Sequence 8 (SEQ ID NO: 2) disclosed below, accommodate the homologies shared among preferred morphogen protein family members identified to date, including at least OP1, OP2, OP3, CBMP2A, CBMP2B, BMP3, BMP5, BMP6, DPP, Vg1, Vgr1, 60A, and GDF1. The amino acid sequences for these proteins are described herein (see Sequence Listing) and/or in the art, as summarized above. The generic sequences include both the amino acid identity shared by these sequences in the C-terminal domain, defined by the six and seven cysteine skeletons (Generic Sequences 7 and 8, respectively), as well as alternative residues for the variable positions within the sequence. The generic sequences provide an appropriate cysteine skeleton where inter- or intramolecular disulfide bonds can form, and contain certain critical amino acids likely to influence the tertiary structure of the folded proteins. In addition, the generic sequences allow for an additional cysteine at position 41 (Generic Sequence 7) or position 46 (Generic Sequence 8), thereby encompassing the morphogenically active sequences of OP2 and OP3.

Generic Sequence 7 (SEQ ID NO: 1)

Leu Xaa Xaa Xaa Phe Xaa Xaa
1 5

Xaa Gly Trp Xaa Xaa Xaa Xaa Xaa Xaa Pro
10 15

Xaa Xaa Xaa Xaa Ala Xaa Tyr Cys Xaa Gly
20 25

Xaa Cys Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa
30 35

Xaa Xaa Xaa Asn His Ala Xaa Xaa Xaa Xaa
40 45

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
		50					55		
Xaa	Xaa	Xaa	Cys	Cys	Xaa	Pro	Xaa	Xaa	Xaa
		60					65		
Xaa	Xaa	Xaa	Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Xaa
		70					75		
Xaa	Xaa	Xaa	Val	Xaa	Leu	Xaa	Xaa	Xaa	Xaa
		80					85		
Xaa	Met	Xaa	Val	Xaa	Xaa	Cys	Xaa	Cys	Xaa
		90					95		

wherein each Xaa independently is selected from a group of one or more specified amino acids defined as follows: "res." means "residue" and Xaa at res. 2 = (Tyr or Lys); Xaa at res. 3 = Val or Ile); Xaa at res. 4 = (Ser, Asp or Glu); Xaa at res. 6 = (Arg, Gln, Ser, Lys or Ala); Xaa at res. 7 = (Asp or Glu); Xaa at res. 8 = (Leu, Val or Ile); Xaa at res. 11 = (Gln, Leu, Asp, His, Asn or Ser); Xaa at res. 12 = (Asp, Arg, Asn or Glu); Xaa at res. 13 = (Trp or Ser); Xaa at res. 14 = (Ile or Val); Xaa at res. 15 = (Ile or Val); Xaa at res. 16 (Ala or Ser); Xaa at res. 18 = (Glu, Gln, Leu, Lys, Pro or Arg); Xaa at res. 19 = (Gly or Ser); Xaa at res. 20 = (Tyr or Phe); Xaa at res. 21 = (Ala, Ser, Asp, Met, His, Gln, Leu or Gly); Xaa at res. 23 = (Tyr, Asn or Phe); Xaa at res. 26 = (Glu, His, Tyr, Asp, Gln, Ala or Ser); Xaa at res. 28 = (Glu, Lys, Asp, Gln or Ala); Xaa at res. 30 = (Ala, Ser, Pro, Gln, Ile or Asn); Xaa at res. 31 = (Phe, Leu or Tyr); Xaa at res. 33 = (Leu, Val or Met); Xaa at res. 34 = (Asn, Asp, Ala, Thr or Pro); Xaa at res. 35 = (Ser, Asp, Glu, Leu, Ala or Lys); Xaa at res. 36 = (Tyr, Cys, His, Ser or Ile); Xaa at res. 37 = (Met, Phe, Gly or Leu); Xaa at res. 38 = (Asn, Ser or Lys); Xaa at res. 39 = (Ala, Ser, Gly or Pro); Xaa at res. 40 = (Thr, Leu or Ser); Xaa at res. 44 = (Ile, Val or Thr); Xaa at res. 45 = (Val, Leu, Met or Ile); Xaa at res. 46 = (Gln or Arg); Xaa at res. 47 = (Thr, Ala or Ser); Xaa at res. 48 = (Leu or Ile); Xaa at res. 49 = (Val or Met); Xaa at res. 50 = (His, Asn or Arg); Xaa at res. 51 = (Phe, Leu, Asn, Ser, Ala or Val); Xaa at res. 52 = (Ile, Met, Asn, Ala, Val, Gly or Leu); Xaa at res. 53 = (Asn, Lys, Ala, Glu, Gly or Phe); Xaa at res. 54 = (Pro, Ser or Val); Xaa at res. 55 = (Glu, Asp, Asn, Gly, Val, Pro or Lys); Xaa at res. 56 = (Thr, Ala, Val, Lys, Asp, Tyr, Ser, Gly, Ile or His); Xaa at res. 57 = (Val, Ala or Ile); Xaa at res. 58 = (Pro or Asp); Xaa at res. 59 = (Lys, Leu or Glu); Xaa at res. 60 = (Pro, Val or Ala); Xaa at res. 63 = (Ala or Val); Xaa at res. 65 = (Thr, Ala or Glu); Xaa at res. 66 = (Gln, Lys, Arg or Glu); Xaa at res. 67 = (Leu, Met or Val); Xaa at res. 68 =

(Asn, Ser, Asp or Gly); Xaa at res. 69 = (Ala, Pro or Ser); Xaa at res. 70 = (Ile, Thr, Val or Leu); Xaa at res. 71 = (Ser, Ala or Pro); Xaa at res. 72 = (Val, Leu, Met or Ile); Xaa at res. 74 = (Tyr or Phe); Xaa at res. 75 = (Phe, Tyr, Leu or His); Xaa at res. 76 = (Asp, Asn or Leu); Xaa at res. 77 = (Asp, Glu, Asn, Arg or Ser); Xaa at res. 78 = (Ser, Gln, Asn, Tyr or Asp); Xaa at res. 79 = (Ser, Asn, Asp, Glu or Lys); Xaa at res. 80 = (Asn, Thr or Lys); Xaa at res. 82 = (Ile, Val or Asn); Xaa at res. 84 = (Lys or Arg); Xaa at res. 85 = (Lys, Asn, Gln, His, Arg or Val); Xaa at res. 86 = (Tyr, Glu or His); Xaa at res. 87 = (Arg, Gln, Glu or Pro); Xaa at res. 88 = (Asn, Glu, Trp or Asp); Xaa at res. 90 = (Val, Thr, Ala or Ile); Xaa at res. 92 = (Arg, Lys, Val, Asp, Gln or Glu); Xaa at res. 93 = (Ala, Gly, Glu or Ser); Xaa at res. 95 = (Gly or Ala) and Xaa at res. 97 = (His or Arg).

Generic Sequence 8 (SEQ ID NO: 2) includes all of Generic Sequence 7 and in addition includes the following sequence (SEQ ID NO: 14) at its N-terminus:

```
Cys   Xaa   Xaa   Xaa   Xaa
 1                        5
```

Accordingly, beginning with residue 7, each "Xaa" in Generic Sequence 8 is a specified amino acid defined as for Generic Sequence 7, with the distinction that each residue number described for Generic Sequence 7 is shifted by five in Generic Sequence 8. Thus, "Xaa at res. 2 =(Tyr or Lys)" in Generic Sequence 7 refers to Xaa at res. 7 in Generic Sequence 8. In Generic Sequence 8, Xaa at res. 2 = (Lys, Arg, Ala or Gln); Xaa at res. 3 = (Lys, Arg or Met); Xaa at res. 4 = (His, Arg or Gln); and Xaa at res. 5 = (Glu, Ser, His, Gly, Arg, Pro, Thr, or Tyr).

As noted above, certain currently preferred morphogen polypeptide sequences useful in this invention have greater than 60% identity, preferably greater than 65% identity, with the amino acid sequence defining the conserved six or seven cysteine skeleton of hOP1 (e.g., residues 43-139 or 38-139 of SEQ ID NO: 4). These particularly preferred sequences include allelic and phylogenetic counterpart variants of the OP1 and OP2 proteins, including the Drosophila 60A protein (SEQ ID NO: 24). Accordingly, in certain particularly preferred embodiments, useful morphogens include active proteins comprising pairs of polypeptide chains within the generic

amino acid sequence herein referred to as "OPX" (SEQ ID NO: 3), which corresponds to the seven cysteine skeleton and accommodates the homologies between several identified variants of OP1 and OP2. As described therein, each Xaa at a given position independently is selected from the residues occurring at the corresponding position in the C-terminal sequence of mouse or human OP1 or OP2 (see SEQ ID NOs: 4-7 and/or SEQ ID NOs: 15-22).

In still other preferred embodiments, useful morphogen polypeptides have amino acid sequences comprising a sequence encoded by a nucleic acid that hybridizes, under stringent hybridization conditions, to DNA or RNA encoding reference morphogen sequences, e.g., C-terminal sequences defining the conserved seven cysteine domains of OP1 or OP2, e.g., nucleotides 1036-1341 and nucleotides 1390-1695 of SEQ ID NO: 15 and 19, respectively. As used herein, stringent hybridization conditions are defined as hybridization according to known techniques in 40% formamide, 5 X SSPE, 5 X Denhardt's Solution, and 0.1% SDS at 37°C overnight, and washing in 0.1 X SSPE, 0.1% SDS at 50°C.

As noted above, morphogens useful in the present invention generally are dimeric proteins comprising a folded pair of the above polypeptides. Morphogens are inactive when reduced, but are active as oxidized homodimers and when oxidized in combination with other morphogens of this invention to produce heterodimers. Thus, members of a folded pair of morphogen polypeptides in a morphogenically active protein can be selected independently from any of the specific morphogen polypeptides mentioned above.

The morphogens useful in the methods, compositions and devices of this invention include proteins comprising any of the polypeptide chains described above, whether isolated from naturally-occurring sources, or produced by recombinant DNA or other synthetic techniques, and includes allelic and phylogenetic counterpart variants of these proteins, as well as biosynthetic variants (muteins) thereof, and various truncated and fusion constructs. Deletion or addition mutants also are envisioned to be active, including those which may alter the conserved C-terminal six or seven cysteine domain, provided that the alteration does not functionally disrupt the relationship of these cysteines in the folded, biologically active, structure. Accordingly, such active forms are considered the equivalent of the specifically described constructs disclosed herein. The proteins may include forms having varying glycosylation patterns, varying N-termini, a family of related proteins having regions of amino acid sequence homology, and active truncated

or mutated forms of native or biosynthetic proteins, produced by expression of recombinant DNA in host cells.

The morphogenic proteins can be expressed from intact or truncated cDNA or from synthetic DNAs in prokaryotic or eukaryotic host cells, and purified, cleaved, refolded, and dimerized to form morphogenically active compositions. Currently preferred host cells include E. coli or mammalian cells, such as CHO, COS or BSC cells. A detailed description of the morphogens useful in the methods, compositions and devices of this invention is disclosed in published application WO92/15323, the disclosure of which is incorporated by reference herein.

Thus, in view of this disclosure, skilled genetic engineers can isolate genes from cDNA or genomic libraries of various different biological species, which encode appropriate amino acid sequences, or construct DNAs from oligonucleotides, and then can express them in various types of host cells, including both prokaryotes and eukaryotes, to produce large quantities of active proteins capable of stimulating the morphogenesis of, and/or inhibiting damage or loss of, mammalian myocardial tissue.

As noted above, a protein is morphogenic herein generally if it induces the developmental cascade of cellular and molecular events that culminate in the formation of new, organ-specific tissue. Preferably, a morphogen comprises a pair of polypeptides having a sequence that corresponds to or is functionally equivalent to at least the conserved C-terminal six or seven cysteine skeleton of human OP1, included in SEQ ID NO: 4. The morphogens generally are competent to induce a cascade of events including all of the following, in a morphogenically permissive environment: stimulating proliferation of progenitor cells; stimulating the differentiation of progenitor cells; stimulating the proliferation of differentiated cells; and supporting the growth and maintenance of differentiated cells. Details of how the morphogens useful in this invention first were identified, as well as a description on how to make, use and test them for morphogenic activity are disclosed in published application WO92/15323. As disclosed therein, the morphogens can be purified from naturally-sourced material or recombinantly produced from prokaryotic or eukaryotic host cells, using the genetic sequences disclosed therein. Alternatively, novel morphogenic sequences can be identified following the procedures disclosed therein.

Exemplary useful morphogens include naturally derived proteins comprising a pair of polypeptides, the amino acid sequences of which comprise sequences selected from those disclosed in the Sequence Listing and Figure 1. Other useful sequences include those of the naturally derived morphogens dorsalin-1, SCREW, NODAL, UNIVIN and GDF5, discussed herein in connection with Table 1, as well as biosynthetic constructs disclosed in U.S. Pat. 5,011,691, the disclosure of which is incorporated herein by reference (e.g., COP1, COP3, COP4, COP5, COP7, and COP16).

Accordingly, certain preferred morphogens useful in the methods and compositions of this invention can be described as morphogenically active proteins having amino acid sequences sharing 70% or, preferably, 80% homology with a reference morphogen sequence described above, e.g., residues 38-139 of SEQ ID NO: 4, where "homology" is as defined herein above. Alternatively, in other preferred embodiments, morphogens useful in the methods and compositions disclosed herein fall within the family of polypeptides described by Generic Sequence 7, SEQ ID NO: 1, more preferably by Generic Sequence 8, SEQ ID NO: 2.

Figure 1 herein sets forth an alignment of the amino acid sequences of the active regions of exemplary naturally occurring proteins that have been identified or appreciated herein as morphogens, including human OP1 (hOP1, SEQ ID NOs: 4 and 15-16), mouse OP1 (mOP1, SEQ ID NOs: 5 and 17-18), human and mouse OP2 (SEQ ID NOs: 6, 7, and 19-22), mouse OP3 (SEQ ID NOs: 25-26), CBMP2A (SEQ ID NO: 8), CBMP2B (SEQ ID NO: 9), BMP3 (SEQ ID NO: 27), DPP (from Drosophila, SEQ ID NO: 10), Vgl, (from Xenopus, SEQ ID NO: 11), Vgr1 (from mouse, SEQ ID NO: 12), GDF1 (from mouse and/or human, SEQ ID NOs: 13, 30 and 31), 60A protein (from Drosophila, SEQ ID NOs: 23 and 24), BMP5 (SEQ ID NO: 28) and BMP6 (SEQ ID NO: 29). The sequences are aligned essentially following the method of Needleman, et al. (1970) J. Mol. Biol., 48:443-453, calculated using the Align Program (DNAstar, Inc.). In Figure 1, three dots indicates that the amino acid in that position is the same as the corresponding amino acid in hOP1. Three dashes indicates that no amino acid is present in that position, and are included for purposes of illustrating homologies. For example, amino acid residue 60 is "missing" in both CBMP2A and CBMP2B. Of course, both of these amino acid sequences in this region comprise Asn-Ser (residues 58, 59), with CBMP2A then comprising Lys and Ile, whereas CBMP-2B comprises Ser and Ile. Figure 1 also illustrates the handling of insertions in the morphogen

amino acid sequence: between residues 56 and 57 of BMP3 is an inserted Val residue; between residues 43 and 44 of GDF1 is inserted the amino acid sequence, Gly-Gly-Pro-Pro. Such deviations from the reference morphogen sequence are ignored for purposes of calculating the defined relationship between, e.g., GDF1 and hOP1. As is apparent from the amino acid sequence comparisons set forth in Figure 1, significant amino acid changes can be made from the reference sequence while retaining morphogenic activity. For example, while the GDF1 protein sequence depicted in Figure 1 shares only about 50% amino acid identity with the hOP1 sequence described therein, the GDF1 sequence shares greater than 70% amino acid sequence homology with the hOP1 sequence, where "homology" is as defined above.

In other embodiments, as an alternative to the administration of a morphogenic protein, an effective amount of an agent competent to stimulate or induce increased endogenous morphogen expression in a mammal may be administered by any of the routes described herein. Such an inducer of a morphogen may be provided to a mammal, e.g., by local or systemic administration to the mammal or by direct administration to implanted myogenic precursor cells, or may be provided to auxiliary cells co-cultured with myogenic precursor cells. Methods for identifying and testing inducers (stimulating agents) competent to modulate the level of production of morphogens by a given tissue or cell type are described in detail in published applications WO93/05172 and WO93/05751, the teachings of which are incorporated herein by reference. Briefly, candidate compounds can be identified and tested by incubation in vitro with a test tissue or cells thereof, or a cultured cell line derived therefrom, for a time sufficient to allow the compound to affect the production, i.e., the expression and/or secretion, of a morphogen produced by the cells of that tissue. Suitable tissue, or cultured cells of a suitable tissue, preferably can be selected from renal epithelium, ovarian tissue, fibroblasts, and osteoblasts.

In other embodiments, an agent which acts as an agonist of a morphogen receptor may be administered instead of the morphogen itself. Such an agent may also be referred to an a morphogen "mimic," "mimetic," or "analog." Thus, for example, a small peptide or other molecule which can mimic the activity of a morphogen in binding to and activating the morphogen's receptor may be employed as an equivalent of the morphogen. Preferably the agonist is a full agonist, but partial morphogen receptor agonists may also be advantageously employed. Methods of identifying such agonists are known in the art and include assays for

compounds which induce morphogen-mediated responses (e.g., induction of differentiation of metanephric mesenchyme, induction of endochondral bone formation). For example, methods of identifying morphogen inducers or agonists of morphogen receptors may be found in U.S. Ser. No. 08/478,097 filed June 7, 1995 and U.S. Ser. No. 08/507,598 filed July 26, 1995, the disclosures of which are incorporated herein by reference.

In yet other embodiments, a small molecule morphogenic activator may be used for promoting the migration, proliferation, and/or differentiation of myogenic precursor cells by increasing the level of expression of proteins associated with myocardial phenotype. Exemplary methods comprise introducing a small molecule morphogenic activator that regulates some portion or portions of a morphogen-induced regulatory pathway, resulting in an effective increase in expression or activity of myocardium-specific protein. This may result either from stimulating an increase in the endogenous expression of such protein or from a decrease in the inhibition of normal expression of such protein. For example, a small molecule morphogenic activator may act at the type I or type II morphogen receptor; or at the serine/threonine kinase, or other kinase domains of those receptors. Another target of pathway activation is the Smad proteins, including the monomeric, dimeric (including heteromeric and homomeric complexes) or trimeric forms (including heteromeric and homomeric complexes). The Smads have been characterized, and are known in the art. See, e.g., Baker, et al., Curr. Op. Genet. Develop., 7: 467-473 (1997), incorporated by reference herein.

Alternately, a small molecule morphogenic activator may lead to activation of a transcription factor (for example, the X-protein shown in Figure 2) that causes phenotype-specific gene expression (i.e., expression of protein characteristic of myocardium). A small molecule morphogenic activator may act to facilitate, mimic, or, if desired, prevent any one or several of the following: type I and/or type II receptor binding, phosphorylation of the type I receptor, phosphorylation of the Smad molecules, Smad complex formation, Smad translocation into the nucleus, nuclear accumulation of the Smad complex, or transcription modulation of the Smad complex. Furthermore, a small molecule morphogenic activator may act on Smads or Smad complexes to alter tertiary structure, thereby to facilitate or inhibit interaction of the Smad or Smad complex with a receptor kinase domain, other Smads, DNA binding proteins, or DNA itself.

In a particularly-preferred embodiment, a small molecule morphogenic activator is contacted with myogenic precursor cells in vivo or in vitro, or is administered to a patient, wherein the small molecule morphogenic activator facilitates formation of Smad complexes, particularly complexes comprising molecules of Smad1, Smad2, Smad3, Smad4, Smad5 and/or Smad8 in order to induce myogenic precursor cells to migrate, proliferate and/or differentiate into cells expressing markers of a myocardial tissue phenotype. Also in a preferred embodiment, methods comprise administering a small molecule morphogenic activator composition that activates a serine/threonine kinase domain associated with a morphogen type I or type II receptor, thereby to activate the pathway involved in morphogen-induced gene expression. In another embodiment, methods of the invention comprise activating Smad4 association with Smad1, thereby to induce morphogen-responsive phenotype. Methods of the invention may also facilitate Smad interaction with specific nucleic acids, such as promoters of myocardial tissue phenotype-specific gene expression (i.e., expression of genes for a phenotypic protein; a protein associated with preservation, restoration, or enhancement of phenotype, including a protein which is critical for production of non-protein phenotypic markers, such as characteristic lipids or carbohydrates; a protein associated with performance of a phenotypic function or morphology; or a morphogen). Such interaction may be, for example, in association with a transcription control factor that is capable of binding to a regulatory portion of a gene and, simultaneously, to one or more regulatory proteins such as a Smad complex (See Figure 2).

An exemplary morphogen-activated pathway is shown in Figure 2. Morphogens are ligands for the type I and type II receptors. Following phosphorylation of the type I receptor by the type-II receptor, the type I receptor specifically phosphorylates Smad1 homodimers. The type I receptor also specifically phosphorylates Smad5 homodimers. The homodimers then separate to form, in association with a phosphorylated Smad4 molecule, a phosphorylated heteromeric complex comprising at least a Smad1 and a Smad4. A phosphorylated Smad1/Smad5/Smad4 heterotrimer may alternatively be formed. The heteromeric complex then translocates into the nucleus, and accumulates therein. In the nucleus, the Smad complex binds operative DNA, either alone or in association with a specific DNA binding protein (the X-protein in Figure 2), to initiate DNA transcription. The "X-protein" acts as a DNA-binding protein, binding the Smad heteromeric complex to the DNA. The Smad1, Smad2, Smad3 and Smad5 proteins consist of

conserved amino- and carboxy-terminal domains linked by a region that is more divergent among the Smads. The carboxy-terminal domain has an effector function. The amino-terminal domain interacts physically with the carboxy-terminal domain, inhibiting its effector activity, and contributes to DNA binding. Receptor-mediated phosphorylation of the serine residues at the end of the carboxy-terminal domain relieves the carboxy-terminal domain from the inhibitory action of the amino-terminal domain. Phosphorylated Smad molecules form a heteromeric complex with at least one other specific Smad family molecule. The resulting Smad complex then translocates into and accumulates in the cell nucleus. There, the heteromeric Smad complexes regulate transcriptional responses either alone or by specific interaction with a DNA-binding protein, such as forkhead activin signal transducer-1 (FAST1).

Other intracellular pathways are induced by morphogens, and may be affected in the manner described herein by use of a small molecule morphogenic activator.

In a preferred embodiment, a small molecule morphogenic activator for use in the invention is a compound that affects one or more intracellular pathways that normally are under morphogen regulation. Such small molecule morphogenic activators preferably have the ability to enter the cell and target one or more intracellular pathway components in order to stimulate or inhibit their activity. For example, a small molecule morphogenic activator that promotes Smad complex formation between Smad1, Smad4, and Smad5 will stimulate pathways leading to expression of genes encoding phenotype-specific proteins.

One way in which to identify a candidate small molecule morphogenic activator is to assay for the ability of the candidate to modulate the effective systemic or local concentration of a morphogen. This may be done, for example, by incubating the candidate in a cell culture that produces the morphogen, and assaying the culture for a parameter indicative of a change in the production level of the morphogen according the methods of U.S.S.N. 08/451,953 and/or U.S. 5,650,276, the teachings of each of which are incorporated by reference herein. Alternatively, candidate compounds are screened for their ability to induce phenotype-specific protein production in a cell culture in which morphogen activity is not present. Examples of compositions which may be screened for their effect on the production of morphogens or other phenotype-specific proteins include but are not limited to chemicals, biological response modifiers (e.g., lymphokines, cytokines, hormones, or vitamins), plant extracts, microbial broths and

extracts medium conditioned by eukaryotic cells, body fluids, or tissue extracts. Useful candidate compositions then may be tested for in vivo efficacy in a suitable animal model. These compositions then may be used in vivo to up-regulate morphogen-activated regulatory pathways of phenotype-specific protein expression.

A simple method of determining if a small molecule composition has effected a change in the level of a phenotype-specific protein in cultured cells is provided in co-owned, co-pending patent application, U.S.S.N. 08/451,953, the disclosure of which is incorporated by reference herein. The level of a target phenotype-specific protein in a cell resulting from exposure to a small molecule is measured. Alternatively, a change in the activity or amount of an intracellular pathway component is measured in response to application of a candidate small molecule. Candidates having the desired affect on protein production or pathway regulation are selected for use in methods of the invention. If, for example, a composition up-regulates the production of OP-1 by a kidney cell line, it would then be desirable to test systemic administration of this compound in an animal model to determine if it up-regulates the production of OP-1 in vivo. The level of morphogen in the body may be a result of a wide range of physical conditions, e.g., tissue degeneration such as occurs in diseases including arthritis, emphysema, osteoporosis, kidney diseases, lung diseases, cardiomyopathy, and cirrhosis of the liver. The decrease in level of morphogens in the body may also occur as a result of the normal process of aging. The same strategy is used for compositions affecting intracellular pathway components. A composition selected by these screening methods is then used as a treatment or prophylactic.

An appropriate test cell is any cell comprising DNA defining a morphogen-responsive transcription activating element operatively associated with a reporter gene encoding a detectable phenotype-specific gene product. Such DNA can occur naturally in a test cell or can be a transfected DNA. The induced intracellular effect typically is characteristic of morphogenic biological activity, such as Smad activation, or activation of a cascade of biochemical events, such as described above, or involving, for example, cyclic nucleotides, diacylglycerol, and/or and other indicators of intracellular signal transduction such as activation or suppression of gene expression, including induction of mRNA resulting from gene transcription and/or induction of protein synthesis resulting from translation of mRNA transcripts indicative of tissue morphogenesis. Exemplary morphogen-responsive cells are preferably of mammalian origin and include, but are

not limited to, osteogenic progenitor cells; calvaria-derived cells; osteoblasts; osteoclasts; osteosarcoma cells and cells of hepatic or neural origin. Any such morphogen responsive cell can be a suitable test cell for assessing whether a candidate substance is a small molecule morphogenic activator.

A preferred identification method is carried out by exposing a test cell to at least one candidate substance, and detecting whether such exposure induces expression of the detectable phenotype-specific gene product that is in operative association with the morphogen-responsive transcription activating element. Expression of this gene product indicates that the candidate substance induces a morphogen-mediated biological effect. Skilled artisans can, in light of guidance provided herein, construct a test cell with a responsive element from a morphogen-responsive cell and a reporter gene of choice, using recombinant vectors and transfection techniques well-known in the art. There are numerous well-known reporter genes useful herein. These include, for example, chloramphenicol acetyltransferase (CAT), luciferase, human growth hormone (hGH), beta-galactosidase, and assay systems and reagents which are available through commercial sources. As will be appreciated by skilled artisans, the listed reporter genes represent only a few of the possible reporter genes that can be used herein. Examples of such reporter genes can be found in Ausubel et al., Eds., Current Protocols in Molecular Biology, John Wiley & Sons, New York, (1989). Broadly, any gene that encodes a detectable product, e.g., any product having detectable enzymatic activity or against which a specific antibody can be raised, can be used as a reporter gene in the present identification method.

A currently preferred reporter gene system is the firefly luciferase reporter system. Gould, et al., Anal. Biochem., 7:404-408 (1988), incorporated herein by reference. The luciferase assay is fast and sensitive. In this assay system, a lysate of the test cell is prepared and combined with ATP and the substrate luciferin. The encoded enzyme luciferase catalyzes a rapid, ATP-dependent oxidation of the substrate to generate a light-emitting product. The total light output is measured and is proportional to the amount of luciferase present over a wide range of enzyme concentrations. CAT is another frequently used reporter gene system; a major advantage of this system is that it has been an extensively validated and is widely accepted as a measure of promoter activity. Gorman , et al., Mol. Cell. Biol., 2:1044-1051 (1982), incorporated by reference herein. In this system, test cells are transfected with CAT expression vectors and

incubated with the candidate substance within 2-3 days of the initial transfection. Thereafter, cell extracts are prepared. The extracts are incubated with acetyl CoA and radioactive chloramphenicol. Following the incubation, acetylated chloramphenicol is separated from nonacetylated form by thin layer chromatography. In this assay, the degree of acetylation reflects the CAT gene activity with the particular promoter.

Another suitable reporter gene system is based on immunologic detection of hGH. This system is also quick and easy to use. Selden, et al., Mol. Cell. Biol., 6:3173-3179 (1986), incorporated by reference herein. The hGH system is advantageous in that the expressed hGH polypeptide is assayed in the media, rather than in a cell extract. Thus, this system does not require the destruction of the test cells. It will be appreciated that the principle of this reporter gene system is not limited to hGH but rather adapted for use with any polypeptide for which an antibody of acceptable specificity is available or can be prepared.

A small molecule morphogenic activator composition may up-regulate a morphogen-activated pathway by acting at any one or more point. For example, small molecule morphogenic activator potentiation of the pathway may be initiated at the receptor level. Depending on the pathway, the transmembrane receptors may be type I and/or type II, or may be comprise variations on either type I or type II receptors. For example, OP-1 is capable of activating regulatory pathways comprising at least two variations of both type I and type II receptors (ActR-1 and BMPR-1B, and ActRII and BMPR-II, respectively). A small molecule morphogenic activator may stimulate the pathway by acting as a ligand and binding to any of the receptors, thereby inducing phosphorylation of type I receptors and/or Smad molecules. Similarly, a small molecule morphogenic activator may activate the regulatory pathway at the level of the serine/threonine kinase domain of the receptors, thereby stimulating phosphorylation of type I receptors and/or Smad molecules.

As a further alternative, a small molecule morphogenic activator may activate the regulatory pathway at the level of Smad complex formation. A small molecule morphogenic activator may stimulate the formation of Smad family homodimers, heterodimers, or other homomeric or heteromeric complexes. Furthermore, a small molecule morphogenic activator may activate the pathway by interacting with a Smad molecule or Smad complex, thereby altering its tertiary structure.

Alternatively, or in addition, a small molecule morphogenic activator may activate the regulatory pathway by facilitating translocation of a Smad molecule or Smad complex or accumulation of the Smad molecule or Smad complex within the nucleus of the cell. By acting as a DNA binding protein or a transcriptional activator, a small molecule morphogenic activator may activate the regulatory pathway by increasing transcriptional activity caused by the Smad molecule or Smad complex.

Furthermore, a small molecule morphogenic activator can act to stimulate the regulatory pathway by interfering with an inhibitor of the pathway. For example, Smad6 and Smad7, which are structurally different than Smad1, Smad2, Smad3 and Smad5, act as inhibitors of certain types of desirable phenotype-specific protein expression (e.g., by activating TGF-β to induce scar tissue formation). Smad6 forms a stable association with type I receptors and interferes with the phosphorylation of other Smad proteins, including Smad2 and Smad 1, and their subsequent heteromerization with Smad4. Smad7 also forms a stable association with activated type I receptors and blocks access and phosphorylation of certain Smad molecules, thereby preventing formation of certain Smad heteromeric complexes. Smad7 also inhibits nuclear accumulation of Smad heteromeric complexes. A small molecule morphogenic activator may interfere with the inhibitory activity of these Smad proteins by, for example, tightly binding to either one or both proteins and rendering either protein incapable of stable association with type I receptors, or by competitively binding and stimulating the morphogen-activated transmembrane receptors. Alternatively, a small molecule morphogenic activator may activate the inhibitory effects of these Smads in order to inhibit an undesirable effect (e.g., TGFβ activity).

E. Subjects for Treatment

As a general matter, the methods of the present invention may be utilized for any mammalian subject at risk of, or afflicted with, loss of or damage to myocardium. Mammalian subjects which may be treated according to the methods of the invention include, but are not limited to, human subjects or patients. In addition, however, the invention may be employed in the treatment of domesticated mammals which are maintained as human companions (e.g., dogs, cats, horses), which have significant commercial value (e.g., dairy cows, beef cattle, sporting animals), which have significant scientific value (e.g., captive or free specimens of endangered species), or which otherwise have value. In addition, as a general matter, the subjects for

treatment with the methods of the present invention need not present indications for morphogen treatment other than those associated with loss of or damage to myocardium. That is, the subjects for treatment generally are expected to be otherwise free of indications for morphogen treatment. In some number of cases, however, the subjects may present with other symptoms (e.g., osteoporosis, chronic renal failure) for which morphogen treatment also would be indicated. In such cases, the morphogen treatment should be adjusted accordingly to avoid excessive dosing.

One of ordinary skill in the medical or veterinary arts is trained to recognize subjects at risk of, or afflicted with, loss of or damage to myocardium. In particular, clinical and non-clinical indications, as well as accumulated experience, relating to the presently disclosed and other methods of treatment, are expected to inform the skilled practitioner in deciding whether a given individual is a subject at risk of, or afflicted with, loss of or damage to myocardium and whether any particular treatment is best suited to the subject's needs, including treatment according to the present invention.

As a general matter, a mammalian subject may be regarded as a subject at risk of, or afflicted with, loss of or damage to myocardium if that subject has already been diagnosed as at risk of, or afflicted with, loss of or damage to myocardium. Such subjects include, but are not limited to, those which have already suffered a myocardial infarction, which have suffered a physical trauma to the heart, or which have been diagnosed with congestive heart failure.

E. Formulations and Methods of Treatment

The morphogens, morphogen inducers, agonists of morphogen receptors, or small molecule morphogenic activators of the present invention may be provided to myogenic precursor cells by any suitable means, preferably directly (e.g., in vitro or locally after implantation, as by addition to culture medium, injection or topical administration to a tissue locus) or systemically (e.g., parenterally or orally). Preferably, the morphogen, morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator comprises part of an aqueous, physiologically acceptable solution so that in addition to delivery of the desired agent to the target cells, the solution does not otherwise adversely affect the cells' or subject's electrolyte and/or volume balance. The aqueous medium for the agent thus may comprise normal physiologic saline (e.g., 9.85% NaCl, 0.15M, pH 7-7.4). Such an aqueous solution containing the agent can be made, for example, by dissolving or dispersing the agent in 50% ethanol containing acetonitrile in

0.1% trifluoroacetic acid (TFA) or 0.1% HCl, or equivalent solvents. One volume of the resultant solution then is added, for example, to ten volumes of phosphate buffered saline (PBS), which further may include 0.1-0.2% human serum albumin (HSA). The resultant solution preferably is vortexed extensively.

For systemic administration, the morphogens, morphogen inducers, agonists of morphogen receptors, or small molecule morphogenic activators of the present invention may be administered by any route which is compatible with the particular morphogen, inducer, or agonist employed. Where the agent is to be provided parenterally, such as by intravenous, subcutaneous, intramuscular, intraorbital, ophthalmic, intraventricular, intracranial, intracapsular, intraspinal, intracisternal, intraperitoneal, buccal, rectal, vaginal, intranasal or by aerosol administration, the agent preferably comprises part of an aqueous solution. In addition, administration may be by periodic injections of a bolus of the morphogen, inducer, agonist, or small molecule morphogenic activator, or may be made more continuous by intravenous or intraperitoneal administration from a reservoir which is external (e.g., an i.v. bag) or internal (e.g., a bioerodable implant, or a colony of implanted, morphogen-producing cells).

If desired, a given morphogen or other agent may be made more soluble by association with a suitable molecule. For example, association of the mature morphogen dimer with the pro domain results in the pro form of the morphogen which typically is more soluble or dispersible in physiological solutions than the corresponding mature form. In fact, endogenous morphogens are thought to be transported (e.g., secreted and circulated) in the mammalian body in this form. This soluble form of the protein can be obtained from culture medium of morphogen-secreting mammalian cells, e.g., cells transfected with nucleic acid encoding and competent to express the morphogen. Alternatively, a soluble species can be formulated by complexing the mature dimer (or an active fragment thereof) with a morphogen pro domain or a solubility-enhancing fragment thereof (described more fully below). Another molecule capable of enhancing solubility and particularly useful for oral administrations, is casein. For example, addition of 0.2% casein increases solubility of the mature active form of OP1 by 80%. Other components found in milk and/or various serum proteins also may be useful.

Useful solutions for parenteral administration may be prepared by any of the methods well known in the pharmaceutical art, described, for example, in Remington's Pharmaceutical Sciences

(Gennaro, A., ed.), Mack Pub., 1990. Formulations of the therapeutic agents of the invention may include, for example, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, hydrogenated naphthalenes, and the like. Formulations for direct administration, in particular, may include glycerol and other compositions of high viscosity to help maintain the agent at the desired locus. Biocompatible, preferably bioresorbable, polymers, including, for example, hyaluronic acid, collagen, tricalcium phosphate, polybutyrate, lactide, and glycolide polymers and lactide/glycolide copolymers, may be useful excipients to control the release of the agent in vivo. Other potentially useful parenteral delivery systems for these agents include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation administration contain as excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or oily solutions for administration in the form of nasal drops, or as a gel to be applied intranasally. Formulations for parenteral administration may also include glycocholate for buccal administration, methoxysalicylate for rectal administration, or cutric acid for vaginal administration. Suppositories for rectal administration also may be prepared by mixing the morphogen, inducer, agonist, or small molecule morphogenic activator with a non-irritating excipient such as cocoa butter or other compositions which are solid at room temperature and liquid at body temperatures.

Formulations for local or topical administration to a tissue or skin surface may be prepared by dispersing the morphogen, inducer, agonist or small molecule morphogenic activator with an acceptable carrier such as a lotion, cream, ointment or soap. Particularly useful are carriers capable of forming a film or layer over the skin or tissue to localize application and inhibit removal. For local or topical administration to internal tissue surfaces, the agent may be dispersed in a liquid tissue adhesive or other substance known to enhance adsorption to a tissue surface. For example, hydroxypropylcellulose or fibrinogen/thrombin solutions may be used to advantage. Alternatively, tissue-coating solutions, such as pectin-containing formulations may be used.

Alternatively, the agents described herein may be administered orally. Oral administration of proteins as therapeutics generally is not practiced as most proteins are readily degraded by digestive enzymes and acids in the mammalian digestive system before they can be absorbed into the bloodstream. However, the morphogens described herein typically are acid stable and

protease-resistant (see, for example, U.S. Pat. No. 4,968,590). In addition, at least one morphogen, OP1, has been identified in mammary gland extract, colostrum and 57-day milk. Moreover, the OP1 purified from mammary gland extract is morphogenically active and also is detected in the bloodstream. Maternal administration, via ingested milk, may be a natural delivery route of TGFβ superfamily proteins. Letterio et al. (1994), Science 264:1936-1938, report that TGFβ is present in murine milk, and that radiolabeled TGFβ is absorbed by gastrointestinal mucosa of suckling juveniles. Labeled, ingested TGFβ appears rapidly in intact form in the juveniles' body tissues, including lung, heart and liver. Finally, soluble form morphogen, e.g., mature morphogen associated with the pro domain, is morphogenically active. These findings, as well as those disclosed in the examples below, indicate that oral and parenteral administration are viable means for administering TGFβ superfamily proteins, including the morphogens, to an individual. In addition, while the mature forms of certain morphogens described herein typically are sparingly soluble, the morphogen form found in milk (and mammary gland extract and colostrum) is readily soluble, probably by association of the mature, morphogenically active form with part or all of the pro domain of the intact sequence and/or by association with one or more milk components. Accordingly, the compounds provided herein also may be associated with molecules capable of enhancing their solubility in vitro or in vivo.

The compounds provided herein also may be associated with molecules capable of targeting the morphogen, inducer, agonist or small molecule morphogenic activator to the desired tissue. For example, an antibody, antibody fragment, or other binding protein that interacts specifically with a surface molecule on cells of the desired tissue, may be used. Useful targeting molecules may be designed, for example, using the single chain binding site technology disclosed, for example, in U.S. Pat. No. 5,091,513. Targeting molecules can be covalently or non-covalently associated with the morphogen, inducer, agonist, or small molecule morphogenic activator.

As will be appreciated by one of ordinary skill in the art, the formulated compositions contain therapeutically effective amounts of the morphogen, morphogen inducers, agonists of morphogen receptors, or small molecule morphogenic activators. That is, they contain amounts which provide appropriate concentrations of the agent to the mammalian myogenic precursor cells for a time sufficient to stimulate morphogenesis of new and functional myocardium, and/or to

prevent, inhibit or delay further significant loss of myocardium or decline of myocardial function. As will be appreciated by those skilled in the art, the concentration of the compounds described in a therapeutic composition of the present invention will vary depending upon a number of factors, including the biological efficacy of the selected agent, the chemical characteristics (e.g., hydrophobicity) of the compounds employed, the formulation of the compound excipients, the administration route, and the treatment envisioned, including whether the active ingredient will be administered directly to cells in vitro, directly into a tissue site, or systemically. The preferred dosage to be administered also is likely to depend on such variables such as the condition of the diseased or damaged tissues, and the overall health status of the particular subject.

As a general matter, for systemic administration, daily or weekly dosages of 0.00001-1000 mg of a morphogen are sufficient, with 0.0001-100 mg being preferable, and 0.001 to 10 mg being even more preferable. Alternatively, a daily or weekly dosage of 0.01-1000 μg/kg body weight, more preferably 0.1-100 μg/kg body weight, may be advantageously employed. Dosages are preferably administered continuously, but daily, multi-weekly, weekly or monthly dosages may also be employed. In addition, in order to facilitate frequent infusions, implantation of a semi-permanent stent (e.g., intravenous, intraperitoneal, intracisternal or intracapsular) may be advisable. It should be noted that no obvious morphogen induced pathological lesions arise when mature morphogen (e.g., OP1, 20 mg) is administered daily to normal growing rats for 21 consecutive days. Moreover, 10 mg systemic injections of morphogen (e.g., OP1) injected daily for 10 days into normal newborn mice does not produce any gross abnormalities.

The morphogens, inducers, agonists or small molecule morphogenic activators of the invention may, of course, be administered alone or in combination with other molecules known to be beneficial in the treatment of the conditions described herein. Thus, in other embodiments the present invention provides pharmaceutical compositions in which a morphogen, morphogen inducer, agonist of a morphogen receptor, or small molecule morphogenic activator is combined with other agents which promote or enhance the proliferation and differentiation of myogenic precursor cells into new and functional myocardium. Thus, the present invention provides pharmaceutical compositions comprising a morphogen, or morphogen inducer, or agonist of a morphogen receptor, or small molecule morphogenic activator, in combination with one or more of a "muscle extract," conditioned medium from differentiated myotubes grown in culture, bFGF,

IGF, PDGF, LIF, ACTH, MSH, or G-CSF. In each such composition, the ratios or the morphogenic and mitogenic agents may be adjusted based upon their activities, as disclosed in the literature or as determined through simple experimentation, to provide a therapeutically effective dosage of each compound in a single unit dosage. The morphogenic and mitogenic agents in such a composition each preferably comprise at least about 1%, and more preferably more than 5% or 10%, of the dry weight of the composition. The compositions may, however, include other pharmaceutical carriers and active agents, as described above and, generally, in Remington's Pharmaceutical Sciences (Gennaro, A., ed.), Mack Pub., 1990, and, therefore, the morphogenic and mitogenic agents may each comprise a small fraction of the final weight of the pharmaceutical composition.

Practice of the invention, including additional preferred aspects and embodiments thereof, will be still more fully understood from the following examples, which are presented herein for illustration only and should not be construed as limiting the invention in any way.

Examples

Preparation of Soluble Morphogen Complexes

A currently preferred form of the morphogen useful herein, having improved solubility in aqueous solutions, is a dimeric morphogenic protein comprising at least the C-terminal seven cysteine domain characteristic of the morphogen family, complexed with a peptide comprising a pro region of a member of the morphogen family, or a solubility-enhancing fragment thereof, or an allelic, species or other sequence variant thereof. Preferably, the dimeric morphogenic protein is complexed with two pro region peptides. Also, the dimeric morphogenic protein preferably is noncovalently complexed with the pro region peptides. The pro region peptides preferably comprise at least the N-terminal eighteen amino acids that define the pro domain of a given naturally occurring morphogen, or an allelic or phylogenetic counterpart variant thereof. In other preferred embodiments, peptides defining substantially the full length pro domain are used.

Other soluble forms of morphogens include dimers of the uncleaved pro forms of these proteins, as well as "hemi-dimers" wherein one subunit of the dimer is an uncleaved pro form of

the protein, and the other subunit comprises the mature form of the protein, including truncated forms thereof, preferably noncovalently associated with a cleaved pro domain peptide.

As described above and in published application WO94/03600, the teachings of which are incorporated herein by reference, useful pro domains include the full length pro regions, as well as various truncated forms hereof, particularly truncated forms cleaved at proteolytic Arg-Xaa-Xaa-Arg cleavage sites within the pro domain polypeptide. For example, in OP1, possible pro sequences include sequences defined by residues 30-292 (full length form); 48-292; and 158-292. Soluble OP1 complex stability is best enhanced when the pro region comprises the full length form rather than a truncated form, such as the residues 48-292 truncated form, in that residues 30-47 show sequence homology to the N-terminal portions of other morphogens, and currently are believed to have particular utility in enhancing complex stability for all morphogens. Accordingly, currently preferred pro domains include peptides comprising at least the N-terminal fragment, e.g., amino acid residues 30-47 of a naturally occurring morphogen pro domain, or a biosynthetic variant thereof that retains the solubility and/or stability enhancing properties of the naturally-occurring peptide.

As will be appreciated by those having ordinary skill in the art, useful sequences encoding the pro region can be obtained from genetic sequences encoding known morphogens. Alternatively, chimeric pro regions can be constructed from the sequences of one or more known morphogens. Still another option is to create a synthetic sequence variant of one or more known pro region sequences.

In another preferred aspect, useful pro region peptides include polypeptide chains comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a DNA or RNA sequence encoding at least the N-terminal eighteen amino acids of the pro region sequence for OP1 or OP2, e.g., nucleotides 136-192 and 152-211 of SEQ ID NOs: 15 and 19, respectively.

A. Isolation from conditioned media or body fluid

Morphogens are expressed from mammalian cells as soluble complexes. Typically, however the complex is disassociated during purification, generally by exposure to denaturants often added to the purification solutions, such as detergents, alcohols, organic solvents, chaotropic agents and compounds added to reduce the pH of the solution. Provided below is a

currently preferred protocol for purifying the soluble proteins from conditioned media (or, optionally, a body fluid such as serum, cerebrospinal or peritoneal fluid), under non-denaturing conditions. The method is rapid, reproducible and yields isolated soluble morphogen complexes in substantially pure form.

Soluble morphogen complexes can be isolated from conditioned media using a simple, three step chromatographic protocol performed in the absence of denaturants. The protocol involves running the media (or body fluid) over an affinity column, followed by ion exchange and gel filtration chromatographies. The affinity column described below is a Zn-IMAC column. The present protocol has general applicability to the purification of a variety of morphogens, all of which are anticipated to be isolatable using only minor modifications of the protocol described below. An alternative protocol also envisioned to have utility includes an immunoaffinity column, created using standard procedures and, for example, using antibody specific for a given morphogen pro domain (complexed, for example, to a protein A-conjugated Sepharose column). Protocols for developing immunoaffinity columns are well described in the art (see, for example, Guide to Protein Purification, M. Deutscher, ed., Academic Press, San Diego, 1990, particularly sections VII and XI thereof).

In this study, OP1 was expressed in mammalian (CHO, Chinese hamster ovary) cells as described in the art (see, for example, international application US90/05903 (WO91/05802). The CHO cell conditioned media containing 0.5% FBS was initially purified using Immobilized Metal-Ion Affinity Chromatography (IMAC). The soluble OP1 complex from conditioned media binds very selectively to the Zn-IMAC resin and a high concentration of imidazole (50 mM imidazole, pH 8.0) is required for the effective elution of the bound complex. The Zn-IMAC step separates the soluble OP1 from the bulk of the contaminating serum proteins that elute in the flowthrough and 35 mM imidazole wash fractions. The Zn-IMAC purified soluble OP1 is next applied to an S-Sepharose cation-exchange column equilibrated in 20 mM $NaPO_4$ (pH 7.0) with 50 mM NaCl. This S-Sepharose step serves to further purify and concentrate the soluble OP1 complex in preparation for the following gel filtration step. The protein was applied to a Sephacryl S-200HR column equilibrated in TBS. Using substantially the same protocol, soluble morphogens also can be isolated from one or more body fluids, including serum, cerebrospinal fluid or peritoneal fluid.

IMAC was performed using Chelating-Sepharose (Pharmacia) that had been charged with three column volumes of 0.2 M $ZnSO_4$. The conditioned media was titrated to pH 7.0 and applied directly to the Zn-IMAC resin equilibrated in 20 mM HEPES (pH 7.0) with 500 mM NaCl. The Zn-IMAC resin was loaded with 80 mL of starting conditioned media per mL of resin. After loading, the column was washed with equilibration buffer and most of the contaminating proteins were eluted with 35 mM imidazole (pH 7.0) in equilibration buffer. The soluble OP1 complex then is eluted with 50 mM imidazole (pH 8.0) in 20 mM HEPES and 500 mM NaCl.

The 50 mM imidazole eluate containing the soluble OP1 complex was diluted with nine volumes of 20 mM $NaPO_4$ (pH 7.0) and applied to an S-Sepharose (Pharmacia) column equilibrated in 20 mM $NaPO_4$ (pH 7.0) with 50 mM NaCl. The S-Sepharose resin was loaded with an equivalent of 800 mL of starting conditioned media per mL of resin. After loading, the S-Sepharose column was washed with equilibration buffer and eluted with 100 mM NaCl followed by 300 mM and 500 mM NaCl in 20 mM $NaPO_4$ (pH 7.0). The 300 mM NaCl pool was further purified using gel filtration chromatography. Fifty mls of the 300 mM NaCl eluate was applied to a 5.0 X 90 cm Sephacryl S-200HR (Pharmacia) equilibrated in Tris buffered saline (TBS), 50 mM Tris, 150 mM NaCl (pH 7.4). The column was eluted at a flow rate of 5 mL/minute collecting 10 mL fractions. The apparent molecular mass of the soluble OP1 was determined by comparison to protein molecular weight standards (alcohol dehydrogenase (ADH, 150 kDa), bovine serum albumin (BSA, 68 kDa), carbonic anhydrase (CA, 30 kDa) and cytochrome C (cytC, 12.5 kDa). The purity of the S-200 column fractions was determined by separation on standard 15% polyacrylamide SDS gels stained with Coomassie blue. The identity of the mature OP1 and the pro-domain was determined by N-terminal sequence analysis after separation of the mature OP1 from the pro-domain using standard reverse phase C18 HPLC.

The soluble OP1 complex elutes with an apparent molecular weight of 110 kDa. This agrees well with the predicted composition of the soluble OP1 complex with one mature OP1 dimer (35-36 kDa) associated with two pro-domains (39 kDa each). Purity of the final complex can be verified by running the appropriate fraction in a reduced 15% polyacrylamide gel.

The complex components can be verified by running the complex-containing fraction from the S-200 or S-200HR columns over a reverse phase C18 HPLC column and eluting in an acetonitrile gradient (in 0.1% TFA), using standard procedures. The complex is dissociated by

this step, and the pro domain and mature species elute as separate species. These separate species then can be subjected to N-terminal sequencing using standard procedures (see, for example, Guide to Protein Purification, M. Deutscher, ed., Academic Press, San Diego, 1990, particularly pp. 602-613), and the identity of the isolated 36 kDa, 39 kDa proteins confirmed as mature morphogen and isolated, cleaved pro domain, respectively. N-terminal sequencing of the isolated pro domain from mammalian cell produced OP1 revealed two forms of the pro region, the intact form (beginning at residue 30 of SEQ ID NO: 16) and a truncated form, (beginning at residue 48 of SEQ ID NO: 16.) N-terminal sequencing of the polypeptide subunit of the isolated mature species reveals a range of N-termini for the mature sequence, beginning at residues 293, 300, 313, 315, 316, and 318, of SEQ ID NO: 16, all of which are active, as demonstrated by the standard bone morphogenesis assay set forth in published application WO92/15323 as incorporated herein by reference.

B. In Vitro Soluble Morphogen Complex Formation

As an alternative to purifying soluble complexes from culture media or a body fluid, soluble complexes can be formulated from purified pro domains and mature dimeric species. Successful complex formation apparently requires association of the components under denaturing conditions sufficient to relax the folded structure of these molecules, without affecting disulfide bonds. Preferably, the denaturing conditions mimic the environment of an intracellular vesicle sufficiently such that the cleaved pro domain has an opportunity to associate with the mature dimeric species under relaxed folding conditions. The concentration of denaturant in the solution then is decreased in a controlled, preferably step-wise manner, so as to allow proper refolding of the dimer and pro regions while maintaining the association of the pro domain with the dimer. Useful denaturants include 4-6M urea or guanidine hydrochloride (GuHCl), in buffered solutions of pH 4-10, preferably pH 6-8. The soluble complex then is formed by controlled dialysis or dilution into a solution having a final denaturant concentration of less than 0.1-2M urea or GuHCl, preferably 1-2 M urea of GuHCl, which then preferably can be diluted into a physiological buffer. Protein purification/renaturing procedures and considerations are well described in the art, and details for developing a suitable renaturing protocol readily can be determined by one having ordinary skill in the art. One useful text on the subject is Guide to

Protein Purification, M. Deutscher, ed., Academic Press, San Diego, 1990, particularly section V. Complex formation also may be aided by addition of one or more chaperone proteins.

C. Stability of Soluble Morphogen Complexes

The stability of the highly purified soluble morphogen complex in a physiological buffer, e.g., Tris-buffered saline (TBS) and phosphate-buffered saline (PBS), can be enhanced by any of a number of means. The currently preferred method is by means of a pro region that comprises at least the first 18 amino acids of the pro sequence (e.g., residues 30-47 of SEQ ID NO: 16 for OP-1), and preferably is the full length pro region. Residues 30-47 show sequence homology to the N-terminal portion of other morphogens and are believed to have particular utility in enhancing complex stability for all morphogens. Other useful means for enhancing the stability of soluble morphogen complexes include three classes of additives. These additives include basic amino acids (e.g., L-arginine, lysine and betaine); nonionic detergents (e.g., Tween 80 or NonIdet P-120); and carrier proteins (e.g., serum albumin and casein). Useful concentrations of these additives include 1-100 mM, preferably 10-70 mM, including 50 mM, basic amino acid;, 0.01-1.0%, preferably 0.05-0.2%, including 0.1% (v/v) nonionic detergent;, and 0.01-1.0%, preferably 0.05-0.2%, including 0.1% (w/v) carrier protein.

Equivalents

The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative rather than limiting on the invention described herein. Scope of the invention is thus indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are intended to be embraced therein.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: COHEN, CHARLES M.

(ii) TITLE OF INVENTION: TREATMENT OF MAMMALIAN MYOCARDIUM WITH MORPHOGENICALLY-TREATED MYOGENIC PRECURSOR CELLS

(iii) NUMBER OF SEQUENCES: 31

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: TESTA, HURWITZ & THIBEAULT, LLP
(B) STREET: 125 HIGH STREET
(C) CITY: BOSTON
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: TWOMEY, MICHAEL J
(B) REGISTRATION NUMBER: 38,349
(C) REFERENCE/DOCKET NUMBER: CRP-123

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617/248-7000
(B) TELEFAX: 617/248-7100

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..97
(D) OTHER INFORMATION: /label= Generic-Seq-7
/note= "wherein each Xaa is independently selected from a group of one or more specified amino acids as defined in the specification."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Leu Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Pro Xaa Xaa Xaa Xaa Ala Xaa Tyr Cys Xaa Gly Xaa Cys Xaa Xaa Pro
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn His Ala Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Pro
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Val Xaa Leu Xaa Xaa Xaa Xaa Xaa Met Xaa Val Xaa Xaa Cys Xaa Cys
85 90 95

Xaa

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..102
(D) OTHER INFORMATION: /label= Generic-Seq-8
/note= "wherin each Xaa is independently selected from a group of one or more specified amino acids as defined in the specification."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Cys Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Xaa Ala Xaa Tyr Cys Xaa Gly
20 25 30

Xaa Cys Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn His Ala
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

Xaa Cys Cys Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Val Xaa Leu Xaa Xaa Xaa Xaa Xaa Met Xaa Val
85 90 95

Xaa Xaa Cys Xaa Cys Xaa
100

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..102
(D) OTHER INFORMATION: /label= OPX
/note= "WHEREIN EACH XAA IS INDEPENDENTLY SELECTED FROM A GROUP OF ONE OR MORE SPECIFIED AMINO ACIDS AS DEFINED IN THE SPECIFICATION"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Cys Xaa Xaa His Glu Leu Tyr Val Xaa Phe Xaa Asp Leu Gly Trp Xaa
1 5 10 15

Asp Trp Xaa Ile Ala Pro Xaa Gly Tyr Xaa Ala Tyr Tyr Cys Glu Gly
20 25 30

Glu Cys Xaa Phe Pro Leu Xaa Ser Xaa Met Asn Ala Thr Asn His Ala
35 40 45

Ile Xaa Gln Xaa Leu Val His Xaa Xaa Xaa Pro Xaa Xaa Val Pro Lys
50 55 60

Xaa Cys Cys Ala Pro Thr Xaa Leu Xaa Ala Xaa Ser Val Leu Tyr Xaa
65 70 75 80

Asp Xaa Ser Xaa Asn Val Xaa Leu Xaa Lys Xaa Arg Asn Met Val Val
85 90 95

Xaa Ala Cys Gly Cys His
100

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: HIPPOCAMPUS

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..139

(D) OTHER INFORMATION: /label= hOP1-MATURE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys
1               5                  10                  15

Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser Ser
            20                  25                  30

Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg
        35                  40                  45

Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala
    50                  55                  60

Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn
65                  70                  75                  80

Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro
                85                  90                  95

Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile
            100                 105                 110

Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr
        115                 120                 125

Arg Asn Met Val Val Arg Ala Cys Gly Cys His
    130                 135
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: MURIDAE
(F) TISSUE TYPE: EMBRYO

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..139
(D) OTHER INFORMATION: /label= MOP1-MATURE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Ser Thr Gly Gly Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys
1               5                  10                  15

Asn Gln Glu Ala Leu Arg Met Ala Ser Val Ala Glu Asn Ser Ser Ser
            20                  25                  30

Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg
        35                  40                  45
```

Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala
50 55 60

Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn
65 70 75 80

Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro
85 90 95

Asp Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile
100 105 110

Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr
115 120 125

Arg Asn Met Val Val Arg Ala Cys Gly Cys His
130 135

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: HOMO SAPIENS
(F) TISSUE TYPE: HIPPOCAMPUS

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..139
(D) OTHER INFORMATION: /label= HOP2-MATURE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ala Val Arg Pro Leu Arg Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu
1 5 10 15

Pro Gln Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser
20 25 30

His Gly Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln
35 40 45

Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala
50 55 60

Tyr Tyr Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn
65 70 75 80

Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro
85 90 95

Asn Ala Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr
100 105 110

Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His
115 120 125

Arg Asn Met Val Val Lys Ala Cys Gly Cys His
130 135

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: MURIDAE
(F) TISSUE TYPE: EMBRYO

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..139
(D) OTHER INFORMATION: /label= MOP2-MATURE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ala Ala Arg Pro Leu Lys Arg Arg Gln Pro Lys Lys Thr Asn Glu Leu
1 5 10 15

Pro His Pro Asn Lys Leu Pro Gly Ile Phe Asp Asp Gly His Gly Ser
20 25 30

Arg Gly Arg Glu Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Arg
35 40 45

Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala
50 55 60

Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asp Ser Cys Met Asn
65 70 75 80

Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro
85 90 95

Asp Val Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr
100 105 110

Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His
115 120 125

Arg Asn Met Val Val Lys Ala Cys Gly Cys His
130 135

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: bovinae

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..101
(D) OTHER INFORMATION: /label= CBMP-2A-FX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Cys Lys Arg His Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn
1 5 10 15

Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly
20 25 30

Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala
50 55 60

Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp
65 70 75 80

Glu Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu
85 90 95

Gly Cys Gly Cys Arg
100

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: HOMO SAPIENS
(F) TISSUE TYPE: hippocampus

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..101
(D) OTHER INFORMATION: /label= CBMP-2B-FX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn
1 5 10 15

Asp Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly
20 25 30

Asp Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile Pro Lys Ala
50 55 60

Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp
65 70 75 80

Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met Val Val Glu
85 90 95

Gly Cys Gly Cys Arg
100

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: DROSOPHILA MELANOGASTER

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..101
(D) OTHER INFORMATION: /label= DPP-FX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asp
1 5 10 15

Asp Trp Ile Val Ala Pro Leu Gly Tyr Asp Ala Tyr Tyr Cys His Gly
20 25 30

Lys Cys Pro Phe Pro Leu Ala Asp His Phe Asn Ser Thr Asn His Ala
35 40 45

Val Val Gln Thr Leu Val Asn Asn Asn Asn Pro Gly Lys Val Pro Lys
50 55 60

Ala Cys Cys Val Pro Thr Gln Leu Asp Ser Val Ala Met Leu Tyr Leu
65 70 75 80

Asn Asp Gln Ser Thr Val Val Leu Lys Asn Tyr Gln Glu Met Thr Val
85 90 95

Val Gly Cys Gly Cys Arg
100

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: XENOPUS

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..102
(D) OTHER INFORMATION: /label= VGL-FX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Cys Lys Lys Arg His Leu Tyr Val Glu Phe Lys Asp Val Gly Trp Gln
1 5 10 15

Asn Trp Val Ile Ala Pro Gln Gly Tyr Met Ala Asn Tyr Cys Tyr Gly
20 25 30

Glu Cys Pro Tyr Pro Leu Thr Glu Ile Leu Asn Gly Ser Asn His Ala
35 40 45

Ile Leu Gln Thr Leu Val His Ser Ile Glu Pro Glu Asp Ile Pro Leu
50 55 60

Pro Cys Cys Val Pro Thr Lys Met Ser Pro Ile Ser Met Leu Phe Tyr
65 70 75 80

Asp Asn Asn Asp Asn Val Val Leu Arg His Tyr Glu Asn Met Ala Val
85 90 95

Asp Glu Cys Gly Cys Arg
100

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: MURIDAE

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..102
(D) OTHER INFORMATION: /label= VGR-1-FX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Cys Lys Lys His Glu Leu Tyr Val Ser Phe Gln Asp Val Gly Trp Gln
1 5 10 15

Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly
20 25 30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val His Val Met Asn Pro Glu Tyr Val Pro Lys
50 55 60

Pro Cys Cys Ala Pro Thr Lys Val Asn Ala Ile Ser Val Leu Tyr Phe
65 70 75 80

Asp Asp Asn Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
85 90 95

Arg Ala Cys Gly Cys His
100

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 106 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..106
(D) OTHER INFORMATION: /note= "GDF-1 (fx)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Cys Arg Ala Arg Arg Leu Tyr Val Ser Phe Arg Glu Val Gly Trp His
1 5 10 15

Arg Trp Val Ile Ala Pro Arg Gly Phe Leu Ala Asn Tyr Cys Gln Gly
20 25 30

Gln Cys Ala Leu Pro Val Ala Leu Ser Gly Ser Gly Gly Pro Pro Ala
35 40 45

Leu Asn His Ala Val Leu Arg Ala Leu Met His Ala Ala Ala Pro Gly
50 55 60

Ala Ala Asp Leu Pro Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser
65 70 75 80

Val Leu Phe Phe Asp Asn Ser Asp Asn Val Val Leu Arg Gln Tyr Glu
85 90 95

Asp Met Val Val Asp Glu Cys Gly Cys Arg
100 105

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Cys Xaa Xaa Xaa Xaa
1               5
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1822 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: HOMO SAPIENS
(F) TISSUE TYPE: HIPPOCAMPUS

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 49..1341
(C) IDENTIFICATION METHOD: experimental
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "OP1"
/evidence= EXPERIMENTAL
/standard_name= "OP1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
GGTGCGGGCC CGGAGCCCGG AGCCCGGGTA GCGCGTAGAG CCGGCGCG ATG CAC GTG        57
                                                      Met His Val
                                                        1

CGC TCA CTG CGA GCT GCG GCG CCG CAC AGC TTC GTG GCG CTC TGG GCA        105
Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala
  5                  10                  15

CCC CTG TTC CTG CTG CGC TCC GCC CTG GCC GAC TTC AGC CTG GAC AAC        153
Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn
 20                  25                  30                  35

GAG GTG CAC TCG AGC TTC ATC CAC CGG CGC CTC CGC AGC CAG GAG CGG        201
Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg
                 40                  45                  50

CGG GAG ATG CAG CGC GAG ATC CTC TCC ATT TTG GGC TTG CCC CAC CGC        249
Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg
            55                  60                  65
```

```
CCG CGC CCG CAC CTC CAG GGC AAG CAC AAC TCG GCA CCC ATG TTC ATG        297
Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met
        70                  75                  80

CTG GAC CTG TAC AAC GCC ATG GCG GTG GAG GAG GGC GGC GGG CCC GGC        345
Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly Gly Pro Gly
    85                  90                  95

GGC CAG GGC TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC        393
Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly
100                 105                 110                 115

CCC CCT CTG GCC AGC CTG CAA GAT AGC CAT TTC CTC ACC GAC GCC GAC        441
Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp
                120                 125                 130

ATG GTC ATG AGC TTC GTC AAC CTC GTG GAA CAT GAC AAG GAA TTC TTC        489
Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe
            135                 140                 145

CAC CCA CGC TAC CAC CAT CGA GAG TTC CGG TTT GAT CTT TCC AAG ATC        537
His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile
        150                 155                 160

CCA GAA GGG GAA GCT GTC ACG GCA GCC GAA TTC CGG ATC TAC AAG GAC        585
Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp
    165                 170                 175

TAC ATC CGG GAA CGC TTC GAC AAT GAG ACG TTC CGG ATC AGC GTT TAT        633
Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile Ser Val Tyr
180                 185                 190                 195

CAG GTG CTC CAG GAG CAC TTG GGC AGG GAA TCG GAT CTC TTC CTG CTC        681
Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu Phe Leu Leu
                200                 205                 210

GAC AGC CGT ACC CTC TGG GCC TCG GAG GAG GGC TGG CTG GTG TTT GAC        729
Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp
            215                 220                 225

ATC ACA GCC ACC AGC AAC CAC TGG GTG GTC AAT CCG CGG CAC AAC CTG        777
Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu
        230                 235                 240

GGC CTG CAG CTC TCG GTG GAG ACG CTG GAT GGG CAG AGC ATC AAC CCC        825
Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro
    245                 250                 255

AAG TTG GCG GGC CTG ATT GGG CGG CAC GGG CCC CAG AAC AAG CAG CCC        873
Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro
260                 265                 270                 275

TTC ATG GTG GCT TTC TTC AAG GCC ACG GAG GTC CAC TTC CGC AGC ATC        921
Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe Arg Ser Ile
                280                 285                 290

CGG TCC ACG GGG AGC AAA CAG CGC AGC CAG AAC CGC TCC AAG ACG CCC        969
Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro
            295                 300                 305

AAG AAC CAG GAA GCC CTG CGG ATG GCC AAC GTG GCA GAG AAC AGC AGC       1017
Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser
        310                 315                 320
```

```
AGC GAC CAG AGG CAG GCC TGT AAG AAG CAC GAG CTG TAT GTC AGC TTC      1065
Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe
    325                 330                 335

CGA GAC CTG GGC TGG CAG GAC TGG ATC ATC GCG CCT GAA GGC TAC GCC      1113
Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala
340                 345                 350                 355

GCC TAC TAC TGT GAG GGG GAG TGT GCC TTC CCT CTG AAC TCC TAC ATG      1161
Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met
                360                 365                 370

AAC GCC ACC AAC CAC GCC ATC GTG CAG ACG CTG GTC CAC TTC ATC AAC      1209
Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn
            375                 380                 385

CCG GAA ACG GTG CCC AAG CCC TGC TGT GCG CCC ACG CAG CTC AAT GCC      1257
Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala
        390                 395                 400

ATC TCC GTC CTC TAC TTC GAT GAC AGC TCC AAC GTC ATC CTG AAG AAA      1305
Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys
    405                 410                 415

TAC AGA AAC ATG GTG GTC CGG GCC TGT GGC TGC CAC TAGCTCCTCC           1351
Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420                 425                 430

GAGAATTCAG ACCCTTTGGG GCCAAGTTTT TCTGGATCCT CCATTGCTCG CCTTGGCCAG    1411

GAACCAGCAG ACCAACTGCC TTTTGTGAGA CCTTCCCCTC CCTATCCCCA ACTTTAAAGG    1471

TGTGAGAGTA TTAGGAAACA TGAGCAGCAT ATGGCTTTTG ATCAGTTTTT CAGTGGCAGC    1531

ATCCAATGAA CAAGATCCTA CAAGCTGTGC AGGCAAAACC TAGCAGGAAA AAAAAACAAC    1591

GCATAAAGAA AAATGGCCGG GCCAGGTCAT TGGCTGGGAA GTCTCAGCCA TGCACGGACT    1651

CGTTTCCAGA GGTAATTATG AGCGCCTACC AGCCAGGCCA CCCAGCCGTG GGAGGAAGGG    1711

GGCGTGGCAA GGGGTGGGCA CATTGGTGTC TGTGCGAAAG GAAAATTGAC CCGGAAGTTC    1771

CTGTAATAAA TGTCACAATA AAACGAATGA ATGAAAAAAA AAAAAAAAAA A              1822
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 431 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
  1               5                  10                  15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
            20                  25                  30
```

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
35 40 45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
50 55 60

Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
65 70 75 80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
85 90 95

Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
100 105 110

Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
115 120 125

Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
130 135 140

Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
145 150 155 160

Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile
165 170 175

Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile
180 185 190

Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu
195 200 205

Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu
210 215 220

Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg
225 230 235 240

His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser
245 250 255

Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn
260 265 270

Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe
275 280 285

Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser
290 295 300

Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu
305 310 315 320

Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr
325 330 335

Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu
340 345 350

Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn
355 360 365

```
Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His
    370                 375                 380

Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln
385                 390                 395                 400

Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
                405                 410                 415

Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
            420                 425                 430
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1873 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: MURIDAE
(F) TISSUE TYPE: EMBRYO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 104..1393
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "MOP1"
/note= "MOP1 (CDNA)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
CTGCAGCAAG TGACCTCGGG TCGTGGACCG CTGCCCTGCC CCCTCCGCTG CCACCTGGGG      60

CGGCGCGGGC CCGGTGCCCC GGATCGCGCG TAGAGCCGGC GCG ATG CAC GTG CGC      115
                                                Met His Val Arg
                                                  1

TCG CTG CGC GCT GCG GCG CCA CAC AGC TTC GTG GCG CTC TGG GCG CCT      163
Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala Pro
  5                  10                  15                  20

CTG TTC TTG CTG CGC TCC GCC CTG GCC GAT TTC AGC CTG GAC AAC GAG      211
Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn Glu
                 25                  30                  35

GTG CAC TCC AGC TTC ATC CAC CGG CGC CTC CGC AGC CAG GAG CGG CGG      259
Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg Arg
             40                  45                  50

GAG ATG CAG CGG GAG ATC CTG TCC ATC TTA GGG TTG CCC CAT CGC CCG      307
Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg Pro
         55                  60                  65

CGC CCG CAC CTC CAG GGA AAG CAT AAT TCG GCG CCC ATG TTC ATG TTG      355
```

```
Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met Leu
 70                  75                  80

GAC CTG TAC AAC GCC ATG GCG GTG GAG GAG AGC GGG CCG GAC GGA CAG      403
Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Ser Gly Pro Asp Gly Gln
 85                  90                  95                 100

GGC TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC CCC CCT      451
Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly Pro Pro
                105                 110                 115

TTA GCC AGC CTG CAG GAC AGC CAT TTC CTC ACT GAC GCC GAC ATG GTC      499
Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp Met Val
            120                 125                 130

ATG AGC TTC GTC AAC CTA GTG GAA CAT GAC AAA GAA TTC TTC CAC CCT      547
Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe His Pro
        135                 140                 145

CGA TAC CAC CAT CGG GAG TTC CGG TTT GAT CTT TCC AAG ATC CCC GAG      595
Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile Pro Glu
    150                 155                 160

GGC GAA CGG GTG ACC GCA GCC GAA TTC AGG ATC TAT AAG GAC TAC ATC      643
Gly Glu Arg Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Tyr Ile
165                 170                 175                 180

CGG GAG CGA TTT GAC AAC GAG ACC TTC CAG ATC ACA GTC TAT CAG GTG      691
Arg Glu Arg Phe Asp Asn Glu Thr Phe Gln Ile Thr Val Tyr Gln Val
                185                 190                 195

CTC CAG GAG CAC TCA GGC AGG GAG TCG GAC CTC TTC TTG CTG GAC AGC      739
Leu Gln Glu His Ser Gly Arg Glu Ser Asp Leu Phe Leu Leu Asp Ser
            200                 205                 210

CGC ACC ATC TGG GCT TCT GAG GAG GGC TGG TTG GTG TTT GAT ATC ACA      787
Arg Thr Ile Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp Ile Thr
        215                 220                 225

GCC ACC AGC AAC CAC TGG GTG GTC AAC CCT CGG CAC AAC CTG GGC TTA      835
Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu Gly Leu
    230                 235                 240

CAG CTC TCT GTG GAG ACC CTG GAT GGG CAG AGC ATC AAC CCC AAG TTG      883
Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro Lys Leu
245                 250                 255                 260

GCA GGC CTG ATT GGA CGG CAT GGA CCC CAG AAC AAG CAA CCC TTC ATG      931
Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro Phe Met
                265                 270                 275

GTG GCC TTC TTC AAG GCC ACG GAA GTC CAT CTC CGT AGT ATC CGG TCC      979
Val Ala Phe Phe Lys Ala Thr Glu Val His Leu Arg Ser Ile Arg Ser
            280                 285                 290

ACG GGG GGC AAG CAG CGC AGC CAG AAT CGC TCC AAG ACG CCA AAG AAC     1027
Thr Gly Gly Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys Asn
        295                 300                 305

CAA GAG GCC CTG AGG ATG GCC AGT GTG GCA GAA AAC AGC AGC AGT GAC     1075
Gln Glu Ala Leu Arg Met Ala Ser Val Ala Glu Asn Ser Ser Ser Asp
    310                 315                 320
```

```
CAG AGG CAG GCC TGC AAG AAA CAT GAG CTG TAC GTC AGC TTC CGA GAC      1123
Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp
325                 330                 335                 340

CTT GGC TGG CAG GAC TGG ATC ATT GCA CCT GAA GGC TAT GCT GCC TAC      1171
Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr
                345                 350                 355

TAC TGT GAG GGA GAG TGC GCC TTC CCT CTG AAC TCC TAC ATG AAC GCC      1219
Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala
            360                 365                 370

ACC AAC CAC GCC ATC GTC CAG ACA CTG GTT CAC TTC ATC AAC CCA GAC      1267
Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Asp
        375                 380                 385

ACA GTA CCC AAG CCC TGC TGT GCG CCC ACC CAG CTC AAC GCC ATC TCT      1315
Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser
    390                 395                 400

GTC CTC TAC TTC GAC GAC AGC TCT AAT GTC ATC CTG AAG AAG TAC AGA      1363
Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg
405                 410                 415                 420

AAC ATG GTG GTC CGG GCC TGT GGC TGC CAC TAGCTCTTCC TGAGACCCTG        1413
Asn Met Val Val Arg Ala Cys Gly Cys His
                425                 430

ACCTTTGCGG GGCCACACCT TTCCAAATCT TCGATGTCTC ACCATCTAAG TCTCTCACTG    1473

CCCACCTTGG CGAGGAGAAC AGACCAACCT CTCCTGAGCC TTCCCTCACC TCCCAACCGG    1533

AAGCATGTAA GGGTTCCAGA AACCTGAGCG TGCAGCAGCT GATGAGCGCC CTTTCCTTCT    1593

GGCACGTGAC GGACAAGATC CTACCAGCTA CCACAGCAAA CGCCTAAGAG CAGGAAAAAT    1653

GTCTGCCAGG AAAGTGTCCA GTGTCCACAT GGCCCCTGGC GCTCTGAGTC TTTGAGGAGT    1713

AATCGCAAGC CTCGTTCAGC TGCAGCAGAA GGAAGGGCTT AGCCAGGGTG GGCGCTGGCG    1773

TCTGTGTTGA AGGGAAACCA AGCAGAAGCC ACTGTAATGA TATGTCACAA TAAAACCCAT    1833

GAATGAAAAA AAAAAAAAAA AAAAAAAAAA AAAAGAATTC                          1873
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 430 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
  1               5                  10                  15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
             20                  25                  30

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
```

```
         35                  40                  45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
     50                  55                  60

Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
 65                  70                  75                  80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Ser Gly
                 85                  90                  95

Pro Asp Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr
            100                 105                 110

Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp
        115                 120                 125

Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu
    130                 135                 140

Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser
145                 150                 155                 160

Lys Ile Pro Glu Gly Glu Arg Val Thr Ala Ala Glu Phe Arg Ile Tyr
                165                 170                 175

Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Gln Ile Thr
            180                 185                 190

Val Tyr Gln Val Leu Gln Glu His Ser Gly Arg Glu Ser Asp Leu Phe
        195                 200                 205

Leu Leu Asp Ser Arg Thr Ile Trp Ala Ser Glu Glu Gly Trp Leu Val
    210                 215                 220

Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His
225                 230                 235                 240

Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile
                245                 250                 255

Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys
            260                 265                 270

Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Leu Arg
        275                 280                 285

Ser Ile Arg Ser Thr Gly Gly Lys Gln Arg Ser Gln Asn Arg Ser Lys
    290                 295                 300

Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Ser Val Ala Glu Asn
305                 310                 315                 320

Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val
                325                 330                 335

Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly
            340                 345                 350

Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser
        355                 360                 365

Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe
```

370 375 380

Ile Asn Pro Asp Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu
385 390 395 400

Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu
405 410 415

Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420 425 430

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: HIPPOCAMPUS

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 490..1696
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "hOP2-PP"
/note= "hOP2 (cDNA)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GGCGCCGGCA GAGCAGGAGT GGCTGGAGGA GCTGTGGTTG GAGCAGGAGG TGGCACGGCA 60

GGGCTGGAGG GCTCCCTATG AGTGGCGGAG ACGGCCCAGG AGGCGCTGGA GCAACAGCTC 120

CCACACCGCA CCAAGCGGTG GCTGCAGGAG CTCGCCCATC GCCCCTGCGC TGCTCGGACC 180

GCGGCCACAG CCGGACTGGC GGGTACGGCG GCGACAGAGG CATTGGCCGA GAGTCCCAGT 240

CCGCAGAGTA GCCCCGGCCT CGAGGCGGTG GCGTCCCGGT CCTCTCCGTC CAGGAGCCAG 300

GACAGGTGTC GCGCGGCGGG GCTCCAGGGA CCGCGCCTGA GGCCGGCTGC CCGCCCGTCC 360

CGCCCCGCCC CGCCGCCCGC CGCCCGCCGA GCCCAGCCTC CTTGCCGTCG GGGCGTCCCC 420

AGGCCCTGGG TCGGCCGCGG AGCCGATGCG CGCCCGCTGA GCGCCCCAGC TGAGCGCCCC 480

CGGCCTGCC ATG ACC GCG CTC CCC GGC CCG CTC TGG CTC CTG GGC CTG 528
Met Thr Ala Leu Pro Gly Pro Leu Trp Leu Leu Gly Leu
1 5 10

GCG CTA TGC GCG CTG GGC GGG GGC GGC CCC GGC CTG CGA CCC CCG CCC 576
Ala Leu Cys Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg Pro Pro Pro
15 20 25

GGC TGT CCC CAG CGA CGT CTG GGC GCG CGC GAG CGC CGG GAC GTG CAG 624
Gly Cys Pro Gln Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Val Gln
30 35 40 45

```
CGC GAG ATC CTG GCG GTG CTC GGG CTG CCT GGG CGG CCC CGG CCC CGC          672
Arg Glu Ile Leu Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg
                50                  55                  60

GCG CCA CCC GCC GCC TCC CGG CTG CCC GCG TCC GCG CCG CTC TTC ATG          720
Ala Pro Pro Ala Ala Ser Arg Leu Pro Ala Ser Ala Pro Leu Phe Met
            65                  70                  75

CTG GAC CTG TAC CAC GCC ATG GCC GGC GAC GAC GAC GAG GAC GGC GCG          768
Leu Asp Leu Tyr His Ala Met Ala Gly Asp Asp Asp Glu Asp Gly Ala
        80                  85                  90

CCC GCG GAG CGG CGC CTG GGC CGC GCC GAC CTG GTC ATG AGC TTC GTT          816
Pro Ala Glu Arg Arg Leu Gly Arg Ala Asp Leu Val Met Ser Phe Val
    95                  100                 105

AAC ATG GTG GAG CGA GAC CGT GCC CTG GGC CAC CAG GAG CCC CAT TGG          864
Asn Met Val Glu Arg Asp Arg Ala Leu Gly His Gln Glu Pro His Trp
110                 115                 120                 125

AAG GAG TTC CGC TTT GAC CTG ACC CAG ATC CCG GCT GGG GAG GCG GTC          912
Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val
                130                 135                 140

ACA GCT GCG GAG TTC CGG ATT TAC AAG GTG CCC AGC ATC CAC CTG CTC          960
Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val Pro Ser Ile His Leu Leu
            145                 150                 155

AAC AGG ACC CTC CAC GTC AGC ATG TTC CAG GTG GTC CAG GAG CAG TCC         1008
Asn Arg Thr Leu His Val Ser Met Phe Gln Val Val Gln Glu Gln Ser
        160                 165                 170

AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT CTT CAG ACG CTC CGA GCT         1056
Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ala
    175                 180                 185

GGA GAC GAG GGC TGG CTG GTG CTG GAT GTC ACA GCA GCC AGT GAC TGC         1104
Gly Asp Glu Gly Trp Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys
190                 195                 200                 205

TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA CTC CGC CTC TAT GTG GAG         1152
Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu
                210                 215                 220

ACT GAG GAC GGG CAC AGC GTG GAT CCT GGC CTG GCC GGC CTG CTG GGT         1200
Thr Glu Asp Gly His Ser Val Asp Pro Gly Leu Ala Gly Leu Leu Gly
            225                 230                 235

CAA CGG GCC CCA CGC TCC CAA CAG CCT TTC GTG GTC ACT TTC TTC AGG         1248
Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val Thr Phe Phe Arg
        240                 245                 250

GCC AGT CCG AGT CCC ATC CGC ACC CCT CGG GCA GTG AGG CCA CTG AGG         1296
Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg Ala Val Arg Pro Leu Arg
    255                 260                 265

AGG AGG CAG CCG AAG AAA AGC AAC GAG CTG CCG CAG GCC AAC CGA CTC         1344
Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln Ala Asn Arg Leu
270                 275                 280                 285

CCA GGG ATC TTT GAT GAC GTC CAC GGC TCC CAC GGC CGG CAG GTC TGC         1392
Pro Gly Ile Phe Asp Asp Val His Gly Ser His Gly Arg Gln Val Cys
                290                 295                 300
```

```
CGT CGG CAC GAG CTC TAC GTC AGC TTC CAG GAC CTC GGC TGG CTG GAC      1440
Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Asp
            305                 310                 315

TGG GTC ATC GCT CCC CAA GGC TAC TCG GCC TAT TAC TGT GAG GGG GAG      1488
Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu
        320                 325                 330

TGC TCC TTC CCA CTG GAC TCC TGC ATG AAT GCC ACC AAC CAC GCC ATC      1536
Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile
    335                 340                 345

CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA AAC GCA GTC CCC AAG GCG      1584
Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys Ala
350                 355                 360                 365

TGC TGT GCA CCC ACC AAG CTG AGC GCC ACC TCT GTG CTC TAC TAT GAC      1632
Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp
                370                 375                 380

AGC AGC AAC AAC GTC ATC CTG CGC AAA CAC CGC AAC ATG GTG GTC AAG      1680
Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn Met Val Val Lys
            385                 390                 395

GCC TGC GGC TGC CAC T GAGTCAGCCC GCCCAGCCCT ACTGCAG                  1723
Ala Cys Gly Cys His
        400
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 402 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
Met Thr Ala Leu Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu Cys
 1               5                  10                  15

Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg Pro Pro Pro Gly Cys Pro
            20                  25                  30

Gln Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Val Gln Arg Glu Ile
        35                  40                  45

Leu Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg Ala Pro Pro
    50                  55                  60

Ala Ala Ser Arg Leu Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu
65                  70                  75                  80

Tyr His Ala Met Ala Gly Asp Asp Asp Glu Asp Gly Ala Pro Ala Glu
                85                  90                  95

Arg Arg Leu Gly Arg Ala Asp Leu Val Met Ser Phe Val Asn Met Val
            100                 105                 110

Glu Arg Asp Arg Ala Leu Gly His Gln Glu Pro His Trp Lys Glu Phe
```

```
        115                 120                 125

Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala
    130                 135                 140

Glu Phe Arg Ile Tyr Lys Val Pro Ser Ile His Leu Leu Asn Arg Thr
145                 150                 155                 160

Leu His Val Ser Met Phe Gln Val Val Gln Glu Gln Ser Asn Arg Glu
                165                 170                 175

Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ala Gly Asp Glu
            180                 185                 190

Gly Trp Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys Trp Leu Leu
        195                 200                 205

Lys Arg His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Glu Asp
    210                 215                 220

Gly His Ser Val Asp Pro Gly Leu Ala Gly Leu Leu Gly Gln Arg Ala
225                 230                 235                 240

Pro Arg Ser Gln Gln Pro Phe Val Val Thr Phe Phe Arg Ala Ser Pro
                245                 250                 255

Ser Pro Ile Arg Thr Pro Arg Ala Val Arg Pro Leu Arg Arg Arg Gln
            260                 265                 270

Pro Lys Lys Ser Asn Glu Leu Pro Gln Ala Asn Arg Leu Pro Gly Ile
        275                 280                 285

Phe Asp Asp Val His Gly Ser His Gly Arg Gln Val Cys Arg Arg His
    290                 295                 300

Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Asp Trp Val Ile
305                 310                 315                 320

Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu Cys Ser Phe
                325                 330                 335

Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu Gln Ser
            340                 345                 350

Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys Ala Cys Cys Ala
        355                 360                 365

Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn
    370                 375                 380

Asn Val Ile Leu Arg Lys His Arg Asn Met Val Val Lys Ala Cys Gly
385                 390                 395                 400

Cys His
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1926 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:
(A) ORGANISM: MURIDAE
(F) TISSUE TYPE: EMBRYO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 93..1289
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "mOP2-PP"
/note= "mOP2 cDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
GCCAGGCACA GGTGCGCCGT CTGGTCCTCC CCGTCTGGCG TCAGCCGAGC CCGACCAGCT      60

ACCAGTGGAT GCGCGCCGGC TGAAAGTCCG AG ATG GCT ATG CGT CCC GGG CCA      113
                                       Met Ala Met Arg Pro Gly Pro
                                         1               5

CTC TGG CTA TTG GGC CTT GCT CTG TGC GCG CTG GGA GGC GGC CAC GGT      161
Leu Trp Leu Leu Gly Leu Ala Leu Cys Ala Leu Gly Gly Gly His Gly
        10                  15                  20

CCG CGT CCC CCG CAC ACC TGT CCC CAG CGT CGC CTG GGA GCG CGC GAG      209
Pro Arg Pro Pro His Thr Cys Pro Gln Arg Arg Leu Gly Ala Arg Glu
    25                  30                  35

CGC CGC GAC ATG CAG CGT GAA ATC CTG GCG GTG CTC GGG CTA CCG GGA      257
Arg Arg Asp Met Gln Arg Glu Ile Leu Ala Val Leu Gly Leu Pro Gly
40                  45                  50                  55

CGG CCC CGA CCC CGT GCA CAA CCC GCC GCT GCC CGG CAG CCA GCG TCC      305
Arg Pro Arg Pro Arg Ala Gln Pro Ala Ala Ala Arg Gln Pro Ala Ser
                60                  65                  70

GCG CCC CTC TTC ATG TTG GAC CTA TAC CAC GCC ATG ACC GAT GAC GAC      353
Ala Pro Leu Phe Met Leu Asp Leu Tyr His Ala Met Thr Asp Asp Asp
            75                  80                  85

GAC GGC GGG CCA CCA CAG GCT CAC TTA GGC CGT GCC GAC CTG GTC ATG      401
Asp Gly Gly Pro Pro Gln Ala His Leu Gly Arg Ala Asp Leu Val Met
        90                  95                  100

AGC TTC GTC AAC ATG GTG GAA CGC GAC CGT ACC CTG GGC TAC CAG GAG      449
Ser Phe Val Asn Met Val Glu Arg Asp Arg Thr Leu Gly Tyr Gln Glu
    105                 110                 115

CCA CAC TGG AAG GAA TTC CAC TTT GAC CTA ACC CAG ATC CCT GCT GGG      497
Pro His Trp Lys Glu Phe His Phe Asp Leu Thr Gln Ile Pro Ala Gly
120                 125                 130                 135

GAG GCT GTC ACA GCT GCT GAG TTC CGG ATC TAC AAA GAA CCC AGC ACC      545
Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Glu Pro Ser Thr
                140                 145                 150

CAC CCG CTC AAC ACA ACC CTC CAC ATC AGC ATG TTC GAA GTG GTC CAA      593
His Pro Leu Asn Thr Thr Leu His Ile Ser Met Phe Glu Val Val Gln
            155                 160                 165

GAG CAC TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT CTT CAG ACG      641
Glu His Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr
```

```
        170                 175                 180

CTC CGA TCT GGG GAC GAG GGC TGG CTG GTG CTG GAC ATC ACA GCA GCC      689
Leu Arg Ser Gly Asp Glu Gly Trp Leu Val Leu Asp Ile Thr Ala Ala
    185                 190                 195

AGT GAC CGA TGG CTG CTG AAC CAT CAC AAG GAC CTG GGA CTC CGC CTC      737
Ser Asp Arg Trp Leu Leu Asn His His Lys Asp Leu Gly Leu Arg Leu
200                 205                 210                 215

TAT GTG GAA ACC GCG GAT GGG CAC AGC ATG GAT CCT GGC CTG GCT GGT      785
Tyr Val Glu Thr Ala Asp Gly His Ser Met Asp Pro Gly Leu Ala Gly
                220                 225                 230

CTG CTT GGA CGA CAA GCA CCA CGC TCC AGA CAG CCT TTC ATG GTA ACC      833
Leu Leu Gly Arg Gln Ala Pro Arg Ser Arg Gln Pro Phe Met Val Thr
            235                 240                 245

TTC TTC AGG GCC AGC CAG AGT CCT GTG CGG GCC CCT CGG GCA GCG AGA      881
Phe Phe Arg Ala Ser Gln Ser Pro Val Arg Ala Pro Arg Ala Ala Arg
        250                 255                 260

CCA CTG AAG AGG AGG CAG CCA AAG AAA ACG AAC GAG CTT CCG CAC CCC      929
Pro Leu Lys Arg Arg Gln Pro Lys Lys Thr Asn Glu Leu Pro His Pro
    265                 270                 275

AAC AAA CTC CCA GGG ATC TTT GAT GAT GGC CAC GGT TCC CGC GGC AGA      977
Asn Lys Leu Pro Gly Ile Phe Asp Asp Gly His Gly Ser Arg Gly Arg
280                 285                 290                 295

GAG GTT TGC CGC AGG CAT GAG CTC TAC GTC AGC TTC CGT GAC CTT GGC     1025
Glu Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly
                300                 305                 310

TGG CTG GAC TGG GTC ATC GCC CCC CAG GGC TAC TCT GCC TAT TAC TGT     1073
Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys
            315                 320                 325

GAG GGG GAG TGT GCT TTC CCA CTG GAC TCC TGT ATG AAC GCC ACC AAC     1121
Glu Gly Glu Cys Ala Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn
        330                 335                 340

CAT GCC ATC TTG CAG TCT CTG GTG CAC CTG ATG AAG CCA GAT GTT GTC     1169
His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asp Val Val
    345                 350                 355

CCC AAG GCA TGC TGT GCA CCC ACC AAA CTG AGT GCC ACC TCT GTG CTG     1217
Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu
360                 365                 370                 375

TAC TAT GAC AGC AGC AAC AAT GTC ATC CTG CGT AAA CAC CGT AAC ATG     1265
Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn Met
                380                 385                 390

GTG GTC AAG GCC TGT GGC TGC CAC TGAGGCCCCG CCCAGCATCC TGCTTCTACT     1319
Val Val Lys Ala Cys Gly Cys His
            395

ACCTTACCAT CTGGCCGGGC CCCTCTCCAG AGGCAGAAAC CCTTCTATGT TATCATAGCT    1379

CAGACAGGGG CAATGGGAGG CCCTTCACTT CCCCTGGCCA CTTCCTGCTA AAATTCTGGT    1439

CTTTCCCAGT TCCTCTGTCC TTCATGGGGT TTCGGGGCTA TCACCCCGCC CTCTCCATCC    1499
```

TCCTACCCCA AGCATAGACT GAATGCACAC AGCATCCCAG AGCTATGCTA ACTGAGAGGT 1559

CTGGGGTCAG CACTGAAGGC CCACATGAGG AAGACTGATC CTTGGCCATC CTCAGCCCAC 1619

AATGGCAAAT TCTGGATGGT CTAAGAAGGC CCTGGAATTC TAAACTAGAT GATCTGGGCT 1679

CTCTGCACCA TTCATTGTGG CAGTTGGGAC ATTTTTAGGT ATAACAGACA CATACACTTA 1739

GATCAATGCA TCGCTGTACT CCTTGAAATC AGAGCTAGCT TGTTAGAAAA AGAATCAGAG 1799

CCAGGTATAG CGGTGCATGT CATTAATCCC AGCGCTAAAG AGACAGAGAC AGGAGAATCT 1859

CTGTGAGTTC AAGGCCACAT AGAAAGAGCC TGTCTCGGGA GCAGGAAAAA AAAAAAAAAC 1919

GGAATTC 1926

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Ala Met Arg Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu Cys
1 5 10 15

Ala Leu Gly Gly Gly His Gly Pro Arg Pro Pro His Thr Cys Pro Gln
20 25 30

Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Met Gln Arg Glu Ile Leu
35 40 45

Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg Ala Gln Pro Ala
50 55 60

Ala Ala Arg Gln Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr
65 70 75 80

His Ala Met Thr Asp Asp Asp Asp Gly Gly Pro Pro Gln Ala His Leu
85 90 95

Gly Arg Ala Asp Leu Val Met Ser Phe Val Asn Met Val Glu Arg Asp
100 105 110

Arg Thr Leu Gly Tyr Gln Glu Pro His Trp Lys Glu Phe His Phe Asp
115 120 125

Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg
130 135 140

Ile Tyr Lys Glu Pro Ser Thr His Pro Leu Asn Thr Thr Leu His Ile
145 150 155 160

Ser Met Phe Glu Val Val Gln Glu His Ser Asn Arg Glu Ser Asp Leu
165 170 175

Phe Phe Leu Asp Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu

```
            180                 185                 190

Val Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn His His
        195                 200                 205

Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Ala Asp Gly His Ser
    210                 215                 220

Met Asp Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro Arg Ser
225                 230                 235                 240

Arg Gln Pro Phe Met Val Thr Phe Phe Arg Ala Ser Gln Ser Pro Val
                245                 250                 255

Arg Ala Pro Arg Ala Ala Arg Pro Leu Lys Arg Arg Gln Pro Lys Lys
            260                 265                 270

Thr Asn Glu Leu Pro His Pro Asn Lys Leu Pro Gly Ile Phe Asp Asp
        275                 280                 285

Gly His Gly Ser Arg Gly Arg Glu Val Cys Arg Arg His Glu Leu Tyr
    290                 295                 300

Val Ser Phe Arg Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln
305                 310                 315                 320

Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asp
                325                 330                 335

Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His
            340                 345                 350

Leu Met Lys Pro Asp Val Val Pro Lys Ala Cys Cys Ala Pro Thr Lys
        355                 360                 365

Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile
    370                 375                 380

Leu Arg Lys His Arg Asn Met Val Val Lys Ala Cys Gly Cys His
385                 390                 395
```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1368 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1368
(D) OTHER INFORMATION: /label= "60A"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
ATG TCG GGA CTG CGA AAC ACC TCG GAG GCC GTT GCA GTG CTC GCC TCC        48
Met Ser Gly Leu Arg Asn Thr Ser Glu Ala Val Ala Val Leu Ala Ser
  1               5                  10                  15
```

```
CTG GGA CTC GGA ATG GTT CTG CTC ATG TTC GTG GCG ACC ACG CCG CCG          96
Leu Gly Leu Gly Met Val Leu Leu Met Phe Val Ala Thr Thr Pro Pro
            20                  25                  30

GCC GTT GAG GCC ACC CAG TCG GGG ATT TAC ATA GAC AAC GGC AAG GAC         144
Ala Val Glu Ala Thr Gln Ser Gly Ile Tyr Ile Asp Asn Gly Lys Asp
        35                  40                  45

CAG ACG ATC ATG CAC AGA GTG CTG AGC GAG GAC GAC AAG CTG GAC GTC         192
Gln Thr Ile Met His Arg Val Leu Ser Glu Asp Asp Lys Leu Asp Val
    50                  55                  60

TCG TAC GAG ATC CTC GAG TTC CTG GGC ATC GCC GAA CGG CCG ACG CAC         240
Ser Tyr Glu Ile Leu Glu Phe Leu Gly Ile Ala Glu Arg Pro Thr His
65                  70                  75                  80

CTG AGC AGC CAC CAG TTG TCG CTG AGG AAG TCG GCT CCC AAG TTC CTG         288
Leu Ser Ser His Gln Leu Ser Leu Arg Lys Ser Ala Pro Lys Phe Leu
                85                  90                  95

CTG GAC GTC TAC CAC CGC ATC ACG GCG GAG GAG GGT CTC AGC GAT CAG         336
Leu Asp Val Tyr His Arg Ile Thr Ala Glu Glu Gly Leu Ser Asp Gln
            100                 105                 110

GAT GAG GAC GAC GAC TAC GAA CGC GGC CAT CGG TCC AGG AGG AGC GCC         384
Asp Glu Asp Asp Asp Tyr Glu Arg Gly His Arg Ser Arg Arg Ser Ala
        115                 120                 125

GAC CTC GAG GAG GAT GAG GGC GAG CAG CAG AAG AAC TTC ATC ACC GAC         432
Asp Leu Glu Glu Asp Glu Gly Glu Gln Gln Lys Asn Phe Ile Thr Asp
    130                 135                 140

CTG GAC AAG CGG GCC ATC GAC GAG AGC GAC ATC ATC ATG ACC TTC CTG         480
Leu Asp Lys Arg Ala Ile Asp Glu Ser Asp Ile Ile Met Thr Phe Leu
145                 150                 155                 160

AAC AAG CGC CAC CAC AAT GTG GAC GAA CTG CGT CAC GAG CAC GGC CGT         528
Asn Lys Arg His His Asn Val Asp Glu Leu Arg His Glu His Gly Arg
                165                 170                 175

CGC CTG TGG TTC GAC GTC TCC AAC GTG CCC AAC GAC AAC TAC CTG GTG         576
Arg Leu Trp Phe Asp Val Ser Asn Val Pro Asn Asp Asn Tyr Leu Val
            180                 185                 190

ATG GCC GAG CTG CGC ATC TAT CAG AAC GCC AAC GAG GGC AAG TGG CTG         624
Met Ala Glu Leu Arg Ile Tyr Gln Asn Ala Asn Glu Gly Lys Trp Leu
        195                 200                 205

ACC GCC AAC AGG GAG TTC ACC ATC ACG GTA TAC GCC ATT GGC ACC GGC         672
Thr Ala Asn Arg Glu Phe Thr Ile Thr Val Tyr Ala Ile Gly Thr Gly
    210                 215                 220

ACG CTG GGC CAG CAC ACC ATG GAG CCG CTG TCC TCG GTG AAC ACC ACC         720
Thr Leu Gly Gln His Thr Met Glu Pro Leu Ser Ser Val Asn Thr Thr
225                 230                 235                 240

GGG GAC TAC GTG GGC TGG TTG GAG CTC AAC GTG ACC GAG GGC CTG CAC         768
Gly Asp Tyr Val Gly Trp Leu Glu Leu Asn Val Thr Glu Gly Leu His
                245                 250                 255

GAG TGG CTG GTC AAG TCG AAG GAC AAT CAT GGC ATC TAC ATT GGA GCA         816
Glu Trp Leu Val Lys Ser Lys Asp Asn His Gly Ile Tyr Ile Gly Ala
```

```
            260                 265                 270

CAC GCT GTC AAC CGA CCC GAC CGC GAG GTG AAG CTG GAC GAC ATT GGA      864
His Ala Val Asn Arg Pro Asp Arg Glu Val Lys Leu Asp Asp Ile Gly
        275                 280                 285

CTG ATC CAC CGC AAG GTG GAC GAC GAG TTC CAG CCC TTC ATG ATC GGC      912
Leu Ile His Arg Lys Val Asp Asp Glu Phe Gln Pro Phe Met Ile Gly
    290                 295                 300

TTC TTC CGC GGA CCG GAG CTG ATC AAG GCG ACG GCC CAC AGC AGC CAC      960
Phe Phe Arg Gly Pro Glu Leu Ile Lys Ala Thr Ala His Ser Ser His
305                 310                 315                 320

CAC AGG AGC AAG CGA AGC GCC AGC CAT CCA CGC AAG CGC AAG AAG TCG     1008
His Arg Ser Lys Arg Ser Ala Ser His Pro Arg Lys Arg Lys Lys Ser
                325                 330                 335

GTG TCG CCC AAC AAC GTG CCG CTG CTG GAA CCG ATG GAG AGC ACG CGC     1056
Val Ser Pro Asn Asn Val Pro Leu Leu Glu Pro Met Glu Ser Thr Arg
            340                 345                 350

AGC TGC CAG ATG CAG ACC CTG TAC ATA GAC TTC AAG GAT CTG GGC TGG     1104
Ser Cys Gln Met Gln Thr Leu Tyr Ile Asp Phe Lys Asp Leu Gly Trp
        355                 360                 365

CAT GAC TGG ATC ATC GCA CCA GAG GGC TAT GGC GCC TTC TAC TGC AGC     1152
His Asp Trp Ile Ile Ala Pro Glu Gly Tyr Gly Ala Phe Tyr Cys Ser
    370                 375                 380

GGC GAG TGC AAT TTC CCG CTC AAT GCG CAC ATG AAC GCC ACG AAC CAT     1200
Gly Glu Cys Asn Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His
385                 390                 395                 400

GCG ATC GTC CAG ACC CTG GTC CAC CTG CTG GAG CCC AAG AAG GTG CCC     1248
Ala Ile Val Gln Thr Leu Val His Leu Leu Glu Pro Lys Lys Val Pro
                405                 410                 415

AAG CCC TGC TGC GCT CCG ACC AGG CTG GGA GCA CTA CCC GTT CTG TAC     1296
Lys Pro Cys Cys Ala Pro Thr Arg Leu Gly Ala Leu Pro Val Leu Tyr
            420                 425                 430

CAC CTG AAC GAC GAG AAT GTG AAC CTG AAA AAG TAT AGA AAC ATG ATT     1344
His Leu Asn Asp Glu Asn Val Asn Leu Lys Lys Tyr Arg Asn Met Ile
        435                 440                 445

GTG AAA TCC TGC GGG TGC CAT TGA                                    1368
Val Lys Ser Cys Gly Cys His
    450                 455
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 455 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Ser Gly Leu Arg Asn Thr Ser Glu Ala Val Ala Val Leu Ala Ser

1 5 10 15

Leu Gly Leu Gly Met Val Leu Leu Met Phe Val Ala Thr Thr Pro Pro
20 25 30

Ala Val Glu Ala Thr Gln Ser Gly Ile Tyr Ile Asp Asn Gly Lys Asp
35 40 45

Gln Thr Ile Met His Arg Val Leu Ser Glu Asp Asp Lys Leu Asp Val
50 55 60

Ser Tyr Glu Ile Leu Glu Phe Leu Gly Ile Ala Glu Arg Pro Thr His
65 70 75 80

Leu Ser Ser His Gln Leu Ser Leu Arg Lys Ser Ala Pro Lys Phe Leu
85 90 95

Leu Asp Val Tyr His Arg Ile Thr Ala Glu Glu Gly Leu Ser Asp Gln
100 105 110

Asp Glu Asp Asp Asp Tyr Glu Arg Gly His Arg Ser Arg Arg Ser Ala
115 120 125

Asp Leu Glu Glu Asp Glu Gly Glu Gln Gln Lys Asn Phe Ile Thr Asp
130 135 140

Leu Asp Lys Arg Ala Ile Asp Glu Ser Asp Ile Ile Met Thr Phe Leu
145 150 155 160

Asn Lys Arg His His Asn Val Asp Glu Leu Arg His Glu His Gly Arg
165 170 175

Arg Leu Trp Phe Asp Val Ser Asn Val Pro Asn Asp Asn Tyr Leu Val
180 185 190

Met Ala Glu Leu Arg Ile Tyr Gln Asn Ala Asn Glu Gly Lys Trp Leu
195 200 205

Thr Ala Asn Arg Glu Phe Thr Ile Thr Val Tyr Ala Ile Gly Thr Gly
210 215 220

Thr Leu Gly Gln His Thr Met Glu Pro Leu Ser Ser Val Asn Thr Thr
225 230 235 240

Gly Asp Tyr Val Gly Trp Leu Glu Leu Asn Val Thr Glu Gly Leu His
245 250 255

Glu Trp Leu Val Lys Ser Lys Asp Asn His Gly Ile Tyr Ile Gly Ala
260 265 270

His Ala Val Asn Arg Pro Asp Arg Glu Val Lys Leu Asp Asp Ile Gly
275 280 285

Leu Ile His Arg Lys Val Asp Asp Glu Phe Gln Pro Phe Met Ile Gly
290 295 300

Phe Phe Arg Gly Pro Glu Leu Ile Lys Ala Thr Ala His Ser Ser His
305 310 315 320

His Arg Ser Lys Arg Ser Ala Ser His Pro Arg Lys Arg Lys Lys Ser
325 330 335

Val Ser Pro Asn Asn Val Pro Leu Leu Glu Pro Met Glu Ser Thr Arg

```
          340                 345                 350

Ser Cys Gln Met Gln Thr Leu Tyr Ile Asp Phe Lys Asp Leu Gly Trp
        355                 360                 365

His Asp Trp Ile Ile Ala Pro Glu Gly Tyr Gly Ala Phe Tyr Cys Ser
    370                 375                 380

Gly Glu Cys Asn Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His
385                 390                 395                 400

Ala Ile Val Gln Thr Leu Val His Leu Leu Glu Pro Lys Lys Val Pro
                405                 410                 415

Lys Pro Cys Cys Ala Pro Thr Arg Leu Gly Ala Leu Pro Val Leu Tyr
            420                 425                 430

His Leu Asn Asp Glu Asn Val Asn Leu Lys Lys Tyr Arg Asn Met Ile
        435                 440                 445

Val Lys Ser Cys Gly Cys His
    450                 455
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1674 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 69..1268
(D) OTHER INFORMATION: /note= "mOP3-PP"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
GGATCCGCGG CGCTGTCCCA TCCTTGTCGT CGAGGCGTCG CTGGATGCGA GTCCGCTAAA      60

CGTCCGAG ATG GCT GCG CGT CCG GGA CTC CTA TGG CTA CTG GGC CTG GCT      110
         Met Ala Ala Arg Pro Gly Leu Leu Trp Leu Leu Gly Leu Ala
           1               5                  10

CTG TGC GTG TTG GGC GGC GGT CAC CTC TCG CAT CCC CCG CAC GTC TTT      158
Leu Cys Val Leu Gly Gly Gly His Leu Ser His Pro Pro His Val Phe
 15                  20                  25                  30

CCC CAG CGT CGA CTA GGA GTA CGC GAG CCC CGC GAC ATG CAG CGC GAG      206
Pro Gln Arg Arg Leu Gly Val Arg Glu Pro Arg Asp Met Gln Arg Glu
                35                  40                  45

ATT CGG GAG GTG CTG GGG CTA GCC GGG CGG CCC CGA TCC CGA GCA CCG      254
Ile Arg Glu Val Leu Gly Leu Ala Gly Arg Pro Arg Ser Arg Ala Pro
            50                  55                  60

GTC GGG GCT GCC CAG CAG CCA GCG TCT GCG CCC CTC TTT ATG TTG GAC      302
Val Gly Ala Ala Gln Gln Pro Ala Ser Ala Pro Leu Phe Met Leu Asp
```

```
          65                  70                  75

CTG TAC CGT GCC ATG ACG GAT GAC AGT GGC GGT GGG ACC CCG CAG CCT      350
Leu Tyr Arg Ala Met Thr Asp Asp Ser Gly Gly Gly Thr Pro Gln Pro
     80                  85                  90

CAC TTG GAC CGT GCT GAC CTG ATT ATG AGC TTT GTC AAC ATA GTG GAA      398
His Leu Asp Arg Ala Asp Leu Ile Met Ser Phe Val Asn Ile Val Glu
 95                 100                 105                 110

CGC GAC CGT ACC CTG GGC TAC CAG GAG CCA CAC TGG AAG GAA TTC CAC      446
Arg Asp Arg Thr Leu Gly Tyr Gln Glu Pro His Trp Lys Glu Phe His
                115                 120                 125

TTT GAC CTA ACC CAG ATC CCT GCT GGG GAG GCT GTC ACA GCT GCT GAG      494
Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala Glu
            130                 135                 140

TTC CGG ATC TAC AAA GAA CCC AGT ACC CAC CCG CTC AAC ACA ACC CTC      542
Phe Arg Ile Tyr Lys Glu Pro Ser Thr His Pro Leu Asn Thr Thr Leu
        145                 150                 155

CAC ATC AGC ATG TTC GAA GTG GTC CAA GAG CAC TCC AAC AGG GAG TCT      590
His Ile Ser Met Phe Glu Val Val Gln Glu His Ser Asn Arg Glu Ser
    160                 165                 170

GAC TTG TTC TTT TTG GAT CTT CAG ACG CTC CGA TCT GGG GAC GAG GGC      638
Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly
175                 180                 185                 190

TGG CTG GTG CTG GAC ATC ACA GCA GCC AGT GAC CGA TGG CTG CTG AAC      686
Trp Leu Val Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn
                195                 200                 205

CAT CAC AAG GAC CTA GGA CTC CGC CTC TAT GTG GAA ACC GAG GAT GGG      734
His His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Glu Asp Gly
            210                 215                 220

CAC AGC ATA GAT CCT GGC CTA GCT GGT CTG CTT GGA CGA CAA GCA CCA      782
His Ser Ile Asp Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro
        225                 230                 235

CGC TCC AGA CAG CCT TTC ATG GTT GGT TTC TTC AGG GCC AAC CAG AGT      830
Arg Ser Arg Gln Pro Phe Met Val Gly Phe Phe Arg Ala Asn Gln Ser
    240                 245                 250

CCT GTG CGG GCC CCT CGA ACA GCA AGA CCA CTG AAG AAG AAG CAG CTA      878
Pro Val Arg Ala Pro Arg Thr Ala Arg Pro Leu Lys Lys Lys Gln Leu
255                 260                 265                 270

AAT CAA ATC AAC CAG CTG CCG CAC TCC AAC AAA CAC CTA GGA ATC CTT      926
Asn Gln Ile Asn Gln Leu Pro His Ser Asn Lys His Leu Gly Ile Leu
                275                 280                 285

GAT GAT GGC CAC GGT TCT CAC GGC AGA GAA GTT TGC CGC AGG CAT GAG      974
Asp Asp Gly His Gly Ser His Gly Arg Glu Val Cys Arg Arg His Glu
            290                 295                 300

CTC TAT GTC AGC TTC CGT GAC CTT GGC TGG CTG GAC TCT GTC ATT GCC     1022
Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Leu Asp Ser Val Ile Ala
        305                 310                 315

CCC CAG GGC TAC TCC GCC TAT TAC TGT GCT GGG GAG TGC ATC TAC CCA     1070
```

```
Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Ala Gly Glu Cys Ile Tyr Pro
    320                 325                 330

CTG AAC TCC TGT ATG AAC TCC ACC AAC CAC GCC ACT ATG CAG GCC CTG      1118
Leu Asn Ser Cys Met Asn Ser Thr Asn His Ala Thr Met Gln Ala Leu
335                 340                 345                 350

GTA CAT CTG ATG AAG CCA GAT ATC ATC CCC AAG GTG TGC TGT GTG CCT      1166
Val His Leu Met Lys Pro Asp Ile Ile Pro Lys Val Cys Cys Val Pro
                355                 360                 365

ACT GAG CTG AGT GCC ATT TCT CTG CTC TAC TAT GAT AGA AAC AAT AAT      1214
Thr Glu Leu Ser Ala Ile Ser Leu Leu Tyr Tyr Asp Arg Asn Asn Asn
            370                 375                 380

GTC ATC CTG CGC AGG GAG CGC AAC ATG GTA GTC CAG GCC TGT GGC TGC      1262
Val Ile Leu Arg Arg Glu Arg Asn Met Val Val Gln Ala Cys Gly Cys
        385                 390                 395

CAC TGAGTCCCTG CCCAACAGCC TGCTGCCATC CCATCTATCT AGTCAGGCCT           1315
His
    400

CTCTTCCAAG GCAGGAAACC AACAAAGAGG GAAGGCAGTG CTTTCAACTC CATGTCCACA   1375

TTCACAGTCT TGGCCCTCTC TGTTCTTTTT GCCAAGGCTG AGAAGATGGT CCTAGTTATA   1435

ACCCTGGTGA CCTCAGTAGC CCGATCTCTC ATCTCCCCAA ACTCCCCAAT GCAGCCAGGG   1495

GCATCTATGT CCTTTGGGAT TGGGCACAGA AGTCCAATTT ACCAACTTAT TCATGAGTCA   1555

CTACTGGCCC AGCCTGGACT TGAACCTGGA ACACAGGGTA GAGCTCAGGC TCTTCAGTAT   1615

CCATCAGAAG ATTTAGGTGT GTGCAGACAT GACCACACTC CCCCTAGCAC TCCATAGCC    1674
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
Met Ala Ala Arg Pro Gly Leu Leu Trp Leu Leu Gly Leu Ala Leu Cys
  1               5                  10                  15

Val Leu Gly Gly Gly His Leu Ser His Pro Pro His Val Phe Pro Gln
             20                  25                  30

Arg Arg Leu Gly Val Arg Glu Pro Arg Asp Met Gln Arg Glu Ile Arg
         35                  40                  45

Glu Val Leu Gly Leu Ala Gly Arg Pro Arg Ser Arg Ala Pro Val Gly
     50                  55                  60

Ala Ala Gln Gln Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr
 65                  70                  75                  80

Arg Ala Met Thr Asp Asp Ser Gly Gly Gly Thr Pro Gln Pro His Leu
```

85 90 95

Asp Arg Ala Asp Leu Ile Met Ser Phe Val Asn Ile Val Glu Arg Asp
100 105 110

Arg Thr Leu Gly Tyr Gln Glu Pro His Trp Lys Glu Phe His Phe Asp
115 120 125

Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg
130 135 140

Ile Tyr Lys Glu Pro Ser Thr His Pro Leu Asn Thr Thr Leu His Ile
145 150 155 160

Ser Met Phe Glu Val Val Gln Glu His Ser Asn Arg Glu Ser Asp Leu
165 170 175

Phe Phe Leu Asp Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu
180 185 190

Val Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn His His
195 200 205

Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Glu Asp Gly His Ser
210 215 220

Ile Asp Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro Arg Ser
225 230 235 240

Arg Gln Pro Phe Met Val Gly Phe Phe Arg Ala Asn Gln Ser Pro Val
245 250 255

Arg Ala Pro Arg Thr Ala Arg Pro Leu Lys Lys Lys Gln Leu Asn Gln
260 265 270

Ile Asn Gln Leu Pro His Ser Asn Lys His Leu Gly Ile Leu Asp Asp
275 280 285

Gly His Gly Ser His Gly Arg Glu Val Cys Arg Arg His Glu Leu Tyr
290 295 300

Val Ser Phe Arg Asp Leu Gly Trp Leu Asp Ser Val Ile Ala Pro Gln
305 310 315 320

Gly Tyr Ser Ala Tyr Tyr Cys Ala Gly Glu Cys Ile Tyr Pro Leu Asn
325 330 335

Ser Cys Met Asn Ser Thr Asn His Ala Thr Met Gln Ala Leu Val His
340 345 350

Leu Met Lys Pro Asp Ile Ile Pro Lys Val Cys Cys Val Pro Thr Glu
355 360 365

Leu Ser Ala Ile Ser Leu Leu Tyr Tyr Asp Arg Asn Asn Asn Val Ile
370 375 380

Leu Arg Arg Glu Arg Asn Met Val Val Gln Ala Cys Gly Cys His
385 390 395

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 104 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..104
(D) OTHER INFORMATION: /note= "BMP3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Cys Ala Arg Arg Tyr Leu Lys Val Asp Phe Ala Asp Ile Gly Trp Ser
1 5 10 15

Glu Trp Ile Ile Ser Pro Lys Ser Phe Asp Ala Tyr Tyr Cys Ser Gly
20 25 30

Ala Cys Gln Phe Pro Met Pro Lys Ser Leu Lys Pro Ser Asn His Ala
35 40 45

Thr Ile Gln Ser Ile Val Ala Arg Ala Val Gly Val Val Pro Gly Ile
50 55 60

Pro Glu Pro Cys Cys Val Pro Glu Lys Met Ser Ser Leu Ser Ile Leu
65 70 75 80

Phe Phe Asp Glu Asn Lys Asn Val Val Leu Lys Val Tyr Pro Asn Met
85 90 95

Thr Val Glu Ser Cys Ala Cys Arg
100

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: HOMO SAPIENS

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..102
(D) OTHER INFORMATION: /note= "BMP5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln
1 5 10 15

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly
20 25 30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val His Leu Met Phe Pro Asp His Val Pro Lys
50 55 60

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
65 70 75 80

Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
85 90 95

Arg Ser Cys Gly Cys His
100

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: HOMO SAPIENS

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..102
(D) OTHER INFORMATION: /note= "BMP6"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Cys Arg Lys His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln
1 5 10 15

Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly
20 25 30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val His Leu Met Asn Pro Glu Tyr Val Pro Lys
50 55 60

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
65 70 75 80

Asp Asp Asn Ser Asn Val Ile Leu Lys Lys Tyr Arg Trp Met Val Val
85 90 95

Arg Ala Cys Gly Cys His
100

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1247 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: HOMO SAPIENS
(F) TISSUE TYPE: BRAIN

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 84..1199
(D) OTHER INFORMATION: /product= "GDF-1"
/note= "GDF-1 CDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
GGGGACACCG GCCCCGCCCT CAGCCCACTG GTCCCGGGCC GCCGCGGACC CTGCGCACTC      60

TCTGGTCATC GCCTGGGAGG AAG ATG CCA CCG CCG CAG CAA GGT CCC TGC          110
                          Met Pro Pro Pro Gln Gln Gly Pro Cys
                            1               5

GGC CAC CAC CTC CTC CTC CTC CTG GCC CTG CTG CTG CCC TCG CTG CCC        158
Gly His His Leu Leu Leu Leu Leu Ala Leu Leu Leu Pro Ser Leu Pro
 10                  15                  20                  25

CTG ACC CGC GCC CCC GTG CCC CCA GGC CCA GCC GCC GCC CTG CTC CAG        206
Leu Thr Arg Ala Pro Val Pro Pro Gly Pro Ala Ala Ala Leu Leu Gln
                 30                  35                  40

GCT CTA GGA CTG CGC GAT GAG CCC CAG GGT GCC CCC AGG CTC CGG CCG        254
Ala Leu Gly Leu Arg Asp Glu Pro Gln Gly Ala Pro Arg Leu Arg Pro
             45                  50                  55

GTT CCC CCG GTC ATG TGG CGC CTG TTT CGA CGC CGG GAC CCC CAG GAG        302
Val Pro Pro Val Met Trp Arg Leu Phe Arg Arg Arg Asp Pro Gln Glu
         60                  65                  70

ACC AGG TCT GGC TCG CGG CGG ACG TCC CCA GGG GTC ACC CTG CAA CCG        350
Thr Arg Ser Gly Ser Arg Arg Thr Ser Pro Gly Val Thr Leu Gln Pro
     75                  80                  85

TGC CAC GTG GAG GAG CTG GGG GTC GCC GGA AAC ATC GTG CGC CAC ATC        398
Cys His Val Glu Glu Leu Gly Val Ala Gly Asn Ile Val Arg His Ile
 90                  95                 100                 105

CCG GAC CGC GGT GCG CCC ACC CGG GCC TCG GAG CCT GTC TCG GCC GCG        446
Pro Asp Arg Gly Ala Pro Thr Arg Ala Ser Glu Pro Val Ser Ala Ala
                110                 115                 120

GGG CAT TGC CCT GAG TGG ACA GTC GTC TTC GAC CTG TCG GCT GTG GAA        494
Gly His Cys Pro Glu Trp Thr Val Val Phe Asp Leu Ser Ala Val Glu
            125                 130                 135

CCC GCT GAG CGC CCG AGC CGG GCC CGC CTG GAG CTG CGT TTC GCG GCG        542
Pro Ala Glu Arg Pro Ser Arg Ala Arg Leu Glu Leu Arg Phe Ala Ala
        140                 145                 150

GCG GCG GCG GCA GCC CCG GAG GGC GGC TGG GAG CTG AGC GTG GCG CAA        590
Ala Ala Ala Ala Ala Pro Glu Gly Gly Trp Glu Leu Ser Val Ala Gln
    155                 160                 165
```

```
GCG GGC CAG GGC GCG GGC GCG GAC CCC GGG CCG GTG CTG CTC CGC CAG        638
Ala Gly Gln Gly Ala Gly Ala Asp Pro Gly Pro Val Leu Leu Arg Gln
170                 175                 180                 185

TTG GTG CCC GCC CTG GGG CCG CCA GTG CGC GCG GAG CTG CTG GGC GCC        686
Leu Val Pro Ala Leu Gly Pro Pro Val Arg Ala Glu Leu Leu Gly Ala
                190                 195                 200

GCT TGG GCT CGC AAC GCC TCA TGG CCG CGC AGC CTC CGC CTG GCG CTG        734
Ala Trp Ala Arg Asn Ala Ser Trp Pro Arg Ser Leu Arg Leu Ala Leu
            205                 210                 215

GCG CTA CGC CCC CGG GCC CCT GCC GCC TGC GCG CGC CTG GCC GAG GCC        782
Ala Leu Arg Pro Arg Ala Pro Ala Ala Cys Ala Arg Leu Ala Glu Ala
        220                 225                 230

TCG CTG CTG CTG GTG ACC CTC GAC CCG CGC CTG TGC CAC CCC CTG GCC        830
Ser Leu Leu Leu Val Thr Leu Asp Pro Arg Leu Cys His Pro Leu Ala
    235                 240                 245

CGG CCG CGG CGC GAC GCC GAA CCC GTG TTG GGC GGC GGC CCC GGG GGC        878
Arg Pro Arg Arg Asp Ala Glu Pro Val Leu Gly Gly Gly Pro Gly Gly
250                 255                 260                 265

GCT TGT CGC GCG CGG CGG CTG TAC GTG AGC TTC CGC GAG GTG GGC TGG        926
Ala Cys Arg Ala Arg Arg Leu Tyr Val Ser Phe Arg Glu Val Gly Trp
                270                 275                 280

CAC CGC TGG GTC ATC GCG CCG CGC GGC TTC CTG GCC AAC TAC TGC CAG        974
His Arg Trp Val Ile Ala Pro Arg Gly Phe Leu Ala Asn Tyr Cys Gln
            285                 290                 295

GGT CAG TGC GCG CTG CCC GTC GCG CTG TCG GGG TCC GGG GGG CCG CCG       1022
Gly Gln Cys Ala Leu Pro Val Ala Leu Ser Gly Ser Gly Gly Pro Pro
        300                 305                 310

GCG CTC AAC CAC GCT GTG CTG CGC GCG CTC ATG CAC GCG GCC GCC CCG       1070
Ala Leu Asn His Ala Val Leu Arg Ala Leu Met His Ala Ala Ala Pro
    315                 320                 325

GGA GCC GCC GAC CTG CCC TGC TGC GTG CCC GCG CGC CTG TCG CCC ATC       1118
Gly Ala Ala Asp Leu Pro Cys Cys Val Pro Ala Arg Leu Ser Pro Ile
330                 335                 340                 345

TCC GTG CTC TTC TTT GAC AAC AGC GAC AAC GTG GTG CTG CGG CAG TAT       1166
Ser Val Leu Phe Phe Asp Asn Ser Asp Asn Val Val Leu Arg Gln Tyr
                350                 355                 360

GAG GAC ATG GTG GTG GAC GAG TGC GGC TGC CGC TAACCCGGGG CGGGCAGGGA      1219
Glu Asp Met Val Val Asp Glu Cys Gly Cys Arg
            365                 370

CCCGGGCCCA ACAATAAATG CCGCGTGG                                        1247
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Pro Pro Pro Gln Gln Gly Pro Cys Gly His His Leu Leu Leu Leu
1 5 10 15

Leu Ala Leu Leu Leu Pro Ser Leu Pro Leu Thr Arg Ala Pro Val Pro
20 25 30

Pro Gly Pro Ala Ala Ala Leu Leu Gln Ala Leu Gly Leu Arg Asp Glu
35 40 45

Pro Gln Gly Ala Pro Arg Leu Arg Pro Val Pro Pro Val Met Trp Arg
50 55 60

Leu Phe Arg Arg Arg Asp Pro Gln Glu Thr Arg Ser Gly Ser Arg Arg
65 70 75 80

Thr Ser Pro Gly Val Thr Leu Gln Pro Cys His Val Glu Glu Leu Gly
85 90 95

Val Ala Gly Asn Ile Val Arg His Ile Pro Asp Arg Gly Ala Pro Thr
100 105 110

Arg Ala Ser Glu Pro Val Ser Ala Ala Gly His Cys Pro Glu Trp Thr
115 120 125

Val Val Phe Asp Leu Ser Ala Val Glu Pro Ala Glu Arg Pro Ser Arg
130 135 140

Ala Arg Leu Glu Leu Arg Phe Ala Ala Ala Ala Ala Ala Ala Pro Glu
145 150 155 160

Gly Gly Trp Glu Leu Ser Val Ala Gln Ala Gly Gln Gly Ala Gly Ala
165 170 175

Asp Pro Gly Pro Val Leu Leu Arg Gln Leu Val Pro Ala Leu Gly Pro
180 185 190

Pro Val Arg Ala Glu Leu Leu Gly Ala Ala Trp Ala Arg Asn Ala Ser
195 200 205

Trp Pro Arg Ser Leu Arg Leu Ala Leu Ala Leu Arg Pro Arg Ala Pro
210 215 220

Ala Ala Cys Ala Arg Leu Ala Glu Ala Ser Leu Leu Leu Val Thr Leu
225 230 235 240

Asp Pro Arg Leu Cys His Pro Leu Ala Arg Pro Arg Arg Asp Ala Glu
245 250 255

Pro Val Leu Gly Gly Gly Pro Gly Gly Ala Cys Arg Ala Arg Arg Leu
260 265 270

Tyr Val Ser Phe Arg Glu Val Gly Trp His Arg Trp Val Ile Ala Pro
275 280 285

Arg Gly Phe Leu Ala Asn Tyr Cys Gln Gly Gln Cys Ala Leu Pro Val
290 295 300

Ala Leu Ser Gly Ser Gly Gly Pro Pro Ala Leu Asn His Ala Val Leu
305 310 315 320

Arg Ala Leu Met His Ala Ala Ala Pro Gly Ala Ala Asp Leu Pro Cys
325 330 335

Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Val Leu Phe Phe Asp Asn
340 345 350

Ser Asp Asn Val Val Leu Arg Gln Tyr Glu Asp Met Val Val Asp Glu
355 360 365

Cys Gly Cys Arg
370

CLAIMS

What is claimed is:

1. A method of therapy for a mammal at risk of, or afflicted with, loss of or damage to myocardium, the method comprising

implanting a preparation of myogenic precursor cells into said mammal at a site at risk of, or afflicted with, loss of or damage to myocardium, and

treating said myogenic precursor cells with an amount of a morphogen sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium.

2. A method of therapy for a mammal at risk of, or afflicted with, loss of or damage to myocardium, the method comprising

implanting a preparation of myogenic precursor cells into said mammal at a site at risk of, or afflicted with, loss of or damage to myocardium, and

treating said mammal with an amount of an inducer of a morphogen encoded by a gene of said mammal, said amount being sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium.

3. A method of therapy for a mammal at risk of, or afflicted with, loss of or damage to myocardium, the method comprising

implanting a preparation of myogenic precursor cells into said mammal at a site at risk of, or afflicted with, loss of or damage to myocardium, and

treating said myogenic precursor cells with an amount of an agonist of a morphogen receptor expressed by said myogenic precursor cells, said amount being sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium.

4. A method of therapy for a mammal at risk of, or afflicted with, loss of or damage to myocardium, the method comprising

implanting a preparation of myogenic precursor cells into said mammal at a site at risk of, or afflicted with, loss of or damage to myocardium, and

treating said myogenic precursor cells with an amount of a small molecule morphogenic activator, said amount being sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium.

5. A method as in any one of claims 1-4 wherein said myogenic precursor cells are selected from the group consisting of mammalian skeletal muscle satellite cells, embryonic myogenic precursor cells, and a histocompatible mammalian myogenic precursor cell line.

6. A method as in any one of claims 1-4 wherein said myogenic precursor cells are autologous skeletal muscle satellite cells.

7. A method as in any one of claims 1-4 wherein

said mammal is afflicted with a condition selected from the group consisting of myocardial infarction and congestive heart failure.

8. A method as in any one of claims 1-4 wherein said treatment step is conducted prior to said implantation step.

9. A method as in any one of claims 1-4 wherein said treatment step is conducted simultaneous with said implantation step.

10. A method as in any one of claims 1-4 wherein said treatment step is conducted subsequent to said implantation step.

11. A method as in claim 10 wherein said treatment step is at least once a week for a period of at least four weeks.

12. A method as in claim 10 wherein said treatment step is at least once a month for a period of at least one year.

13. A method as in claim 1 wherein said morphogen treatment step is conducted with morphogen at a concentration of about 0.01-1000 ng/ml.

14. A method as in claim 1 wherein said morphogen treatment step is conducted with morphogen at a concentration of about 0.1-100 ng/ml.

15. A method of promoting proliferation of myogenic precursor cells or differentiation of myogenic precursor cells into functional myocardium comprising the steps of:

(a) contacting said cells with a morphogen in an amount effective to induce said proliferation or differentiation; and

(b) maintaining said cells in a morphogenically permissive environment.

16. A method as in claim 1 wherein said morphogen is selected from the group consisting of a pro form of a morphogen, a soluble form of a morphogen, a mature morphogen, and a C-terminal fragment of a morphogen comprising at least the seven cysteine domain of said morphogen.

17. A method as in claim 1 wherein said morphogen is selected from the group consisting of osteogenic proteins and bone morphogenic proteins.

18. A method as in claim 1 wherein said morphogen

induces a cascade of tissue-specific morphogenesis culminating in the formation of functional mammalian myocardium; and

comprises a pair of folded polypeptides, the amino acid sequence of each of which comprises a sequence having at least 70% amino acid sequence homology with the C-terminal seven-cysteine domain of human OP-1, mouse OP-1, human OP-2 or mouse OP-2, residues 38-139 of SEQ ID NOs. 5, 6, 7 or 8, respectively.

19. A method as in claim 1 wherein said morphogen is selected from the group consisting of OP-1, CBMP-2A (BMP-2), and CBMP-2B (BMP-4).

20. A therapeutic composition for promoting the repair or regeneration of mammalian myocardium comprising

isolated mammalian myogenic precursor cells, and

an amount of a morphogen sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

21. A therapeutic composition for promoting the repair or regeneration of mammalian myocardium comprising

isolated mammalian myogenic precursor cells, and

an amount of an inducer of a morphogen sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

22. A therapeutic composition for promoting the repair or regeneration of mammalian myocardium comprising

isolated mammalian myogenic precursor cells, and

an amount of an agonist of a morphogen receptor sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

23. A therapeutic composition for promoting the repair or regeneration of mammalian myocardium comprising

isolated mammalian myogenic precursor cells, and

an amount of a small molecule morphogenic activator sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

24. A method of culturing mammalian myogenic precursor cells comprising

isolating said myogenic precursor cells, and

culturing said myogenic precursor cells in a medium comprising an amount of a morphogen sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

25. A method of culturing mammalian myogenic precursor cells comprising

isolating said myogenic precursor cells, and

culturing said myogenic precursor cells in a medium comprising an amount of an inducer of a morphogen sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

26. A method of culturing mammalian myogenic precursor cells comprising

isolating said myogenic precursor cells, and

culturing said myogenic precursor cells in a medium comprising an amount of an agonist of a morphogen receptor sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

27. A method of culturing mammalian myogenic precursor cells comprising

isolating said myogenic precursor cells, and

culturing said myogenic precursor cells in a medium comprising an amount of a small molecule morphogenic activator sufficient to promote proliferation or differentiation of said myogenic precursor cells into functional myocardium in a morphogenically permissive environment.

28. A method of inducing myogenic precursor cells, naturally competent to differentiate into skeletal or smooth muscle, to differentiate into cardiomyocytes, said method comprising the steps of

(a) contacting said myogenic precursor cells with a morphogen; and

(b) maintaining the product of (a) in an environment morphogenically permissive for cardiomyogenesis.

29. A method of producing replacement cardiomyocytes in a mammal in need thereof, said method comprising the step of implanting into said mammal myogenic precursor cells induced by the method of claim 28.

30. A pharmaceutical composition comprising

a morphogenic agent selected from the group consisting of a morphogen, a morphogen inducer, an agonist of a morphogen receptor, and a small molecule morphogenic activator; and

a mitogen selected from the group consisting of bFGF, IGF, PDGF, LIF, ACTH, MSH, and G-CSF.

	Cys	Lys	Lys	His	Glu	Leu	Tyr	Val
hOP-1	...	...	...	...	...	...	...	...
mOP-1	...	...	...	...	...	...	...	...
hOP-2	...	Arg	Arg	...	...	...	...	...
mOP-2	...	Arg	Arg	...	...	...	...	...
mOP-3	...	Arg	Arg	...	...	...	...	...
DPP	...	Arg	Arg	...	Ser	...	...	...
Vg1	...	...	Lys	Arg	His	...	...	...
Vgr-1	...	...	...	...	Gly	...	...	...
CBMP-2A	...	...	Arg	...	Pro	...	...	...
CBMP-2B	...	Arg	Arg	...	Ser	...	...	...
BMP3	...	Ala	Arg	Arg	Tyr	...	Lys	...
GDF-1	...	Arg	Ala	Arg	Arg	...	...	...
60A	...	Gln	Met	Glu	Thr	...	...	...
BMP5	...	...	...	...	...	...	...	...
BMP6	...	Arg	...	...	...	...	...	...
	1				5			

FIG. 1A

```
hOP-1     Ser Phe Arg Asp Leu Gly Trp Gln Asp
mOP-1     ... ... ... ... ... ... ... ... ...
hOP-2     ... ... Gln ... ... ... Leu ... ...
mOP-2     Ser ... ... ... ... ... Leu ... ...
mOP-3     ... ... ... ... ... ... Leu ... ...
DPP       Asp ... Ser ... Val ... ... Asp ...
Vg1       Glu ... Lys ... Val ... ... ... Asn
Vgr-1     ... ... Gln ... Val ... ... ... ...
CBMP-2A   Asp ... Ser ... Val ... ... Asn ...
CBMP-2B   Asp ... Ser ... Val ... ... Asn ...
BMP3      Asp ... Ala ... Ile ... ... Ser Glu
GDF-1     ... ... ... Glu Val ... ... His Arg
60A       Asp ... Lys ... ... ... ... His ...
BMP5      ... ... ... ... ... ... ... ... ...
BMP6      ... ... Gln ... ... ... ... ... ...
                  10                  15
```

FIG. 1B

	Trp	Ile	Ile	Ala	Pro	Glu	Gly	Tyr	Ala
hOP-1	...	...	...	...	...	...	...	...	...
mOP-1	...	...	...	...	...	...	...	...	...
hOP-2	...	Val	...	...	...	Gln	...	...	Ser
mOP-2	...	Val	...	...	...	Gln	...	...	Ser
mOP-3	Ser	Val	...	...	...	Gln	...	...	Ser
DPP	...	...	Val	...	...	Leu	...	...	Asp
Vgl	...	Val	...	...	...	Gln	...	...	Met
Vgr-1	...	...	...	...	...	Lys	...	...	...
CBMP-2A	...	...	Val	...	...	Pro	...	...	His
CBMP-2B	...	...	Val	...	...	Pro	...	...	Gln
BMP3	...	...	...	Ser	...	Lys	Ser	Phe	Asp
GDF-1	...	Val	...	...	...	Arg	...	Phe	Leu
60A	...	...	...	...	...	...	...	...	Gly
BMP5	...	...	...	...	...	...	...	...	...
BMP6	...	...	...	...	...	Lys	...	...	...
			20					25	

FIG. 1C

	Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ala
hOP-1	...	...	...	...	...	...	...	...	...
mOP-1	...	...	...	...	...	...	...	...	...
hOP-2	...	...	...	...	...	...	...	...	Ser
mOP-2	...	...	...	...	...	...	...	...	...
mOP-3	...	...	...	...	Ala	...	...	...	Ile
DPP	...	...	...	...	His	...	Lys	...	Pro
Vg1	...	Asn	...	...	Tyr	...	...	...	Pro
Vgr-1	...	Asn	...	...	Asp	...	...	...	Ser
CBMP-2A	...	Phe	...	...	His	...	Glu	...	Pro
CBMP-2B	...	Phe	...	...	His	...	Asp	...	Pro
BMP3	...	...	...	...	Ser	...	Ala	...	Gln
GDF-1	...	Asn	...	...	Gln	...	Gln	...	...
60A	...	Phe	...	...	Ser	...	...	...	Asn
BMP5	...	Phe	...	...	Asp	...	...	...	Ser
BMP6	...	Asn	...	...	Asp	...	...	...	Ser
				30					35

FIG. 1D

hOP-1	Phe	Pro	Leu	Asn	Ser	Tyr	Met	Asn	Ala
mOP-1	...	...	...	...	...	...	...	...	...
hOP-2	...	...	...	Asp	...	Cys	...	...	...
mOP-2	...	...	...	Asp	...	Cys	...	...	...
mOP-3	Tyr	...	...	...	...	Cys	...	...	Ser
DPP	...	...	...	Ala	Asp	His	Phe	...	Ser
Vgl	Tyr	...	...	Thr	Glu	Ile	Leu	...	Gly
Vgr-1	...	...	...	...	Ala	His	...	...	...
CBMP-2A	...	...	...	Ala	Asp	His	Leu	...	Ser
CBMP-2B	...	...	...	Ala	Asp	His	Leu	...	Ser
GDF-1	Leu	...	Val	Ala	Leu	Ser	Gly	Ser**	...
BMP3	...	...	Met	Pro	Lys	Ser	Leu	Lys	Pro
60A	...	...	...	...	Ala	His	...	...	...
BMP5	...	...	...	...	Ala	His	Met	...	...
BMP6	...	...	...	...	Ala	His	Met	...	...

40

FIG. 1E

```
hOP-1    Thr Asn His Ala Ile Val Gln Thr Leu
mOP-1    ... ... ... ... ... ... ... ... ...
hOP-2    ... ... ... ... Leu ... Ser ... ...
mOP-2    ... ... ... ... Leu ... Ser ... ...
mOP-3    ... ... ... ... Thr Met ... Ala ...
DPP      ... ... ... ... Val ... ... ... ...
Vgl      Ser ... ... ... ... Leu ... ... ...
Vgr-1    ... ... ... ... ... ... ... ... ...
CBMP-2A  ... ... ... ... ... ... ... ... ...
CBMP-2B  ... ... ... ... ... ... ... ... ...
BMP3     Ser ... ... ... Thr Ile ... Ser Ile
GDF-1    Leu ... ... ... Val Leu Arg Ala ...
60A      ... ... ... ... ... ... ... ... ...
BMP5     ... ... ... ... ... ... ... ... ...
BMP6     ... ... ... ... ... ... ... ... ...
         45                  50
```

FIG. 1F

```
hOP-1   Val His Phe Ile Asn Pro Glu Thr Val
mOP-1   ... ... ... ... ... ... ... ... ...
hOP-2   ... His Leu Met Lys ... Asn Ala ...
mOP-2   ... His Leu Met Lys ... Asp Val ...
mOP-3   ... ... Leu Met Lys ... Asp Ile Ile
 DPP    ... Asn Asn Asn ... ... Gly Lys ...
 Vgl    ... ... Ser ... Glu ... ... Asp Ile
Vgr-1   ... ... Val Met ... ... ... Tyr ...
CBMP-2A ... Asn Ser Val ... Ser ... Lys Ile
CBMP-2B ... Asn Ser Val ... Ser ... Ser Ile
 BMP3   ... Arg Ala** Gly Val Val Pro Gly Ile
 GDF-1  Met ... Ala Ala Ala ... Gly Ala Ala
 60A    ... ... Leu Leu Glu ... Lys Lys ...
 BMP5   ... ... Leu Met Phe ... Asp His ...
 BMP6   ... ... Leu Met ... ... ... Tyr ...
                55                  60
```

FIG. 1G

hOP-1	Pro	Lys	Pro	Cys	Cys	Ala	Pro	Thr	Gln	
mOP-1	...	...	...	...	...	...	...	...	...	
hOP-2	...	...	Ala	...	...	...	...	...	Lys	
mOP-2	...	...	Ala	...	...	...	...	...	Lys	
mOP-3	...	...	Val	...	...	Val	...	...	Glu	
DPP	...	...	Ala	...	...	Val	...	...	...	
Vg1	...	Leu	...	...	...	Val	...	...	Lys	
Vgr-1	...	...	...	...	...	...	...	...	Lys	
CBMP-2A	...	...	Ala	...	...	Val	...	...	Glu	
CBMP-2B	...	...	Ala	...	...	Val	...	...	Glu	
BMP3	...	Glu	...	...	...	Val	...	Glu	Lys	
GDF-1	Asp	Leu	...	...	...	Val	...	Ala	Arg	
60A	...	...	...	...	...	...	...	...	Arg	
BMP5	...	...	...	...	...	...	...	...	Lys	
BMP6	...	...	...	...	...	...	...	...	Lys	
			65					70		

FIG. 1H

```
hOP-1   Leu Asn Ala Ile Ser Val Leu Tyr Phe
mOP-1   ... ... ... ... ... ... ... ... ...
hOP-2   ... Ser ... Thr ... ... ... ... Tyr
mOP-2   ... Ser ... Thr ... ... ... ... Tyr
mOP-3   ... Ser ... ... ... Leu ... ... Tyr
  Vgl   Met Ser Pro ... ... Met ... Phe Tyr
 Vgr-1  Val ... ... ... ... ... ... ... ...
  DPP   ... Asp Ser Val Ala Met ... ... Leu
CBMP-2A ... Ser ... ... ... Met ... ... Leu
CBMP-2B ... Ser ... ... ... Met ... ... Leu
 BMP3   Met Ser Ser Leu ... Ile ... Phe Tyr
 GDF-1  ... Ser Pro ... ... ... ... Phe ...
  60A   ... Gly ... Leu Pro ... ... ... His
 BMP5   ... ... ... ... ... ... ... ... ...
 BMP6   ... ... ... ... ... ... ... ... ...
                75                      80
```

FIG. 11

hOP-1	Asp	Asp	Ser	Ser	Asn	Val	Ile	Leu	Lys
mOP-1	...	...	...	...	...	...	...	...	...
hOP-2	...	Ser	...	Asn	...	...	...	...	Arg
mOP-2	...	Ser	...	Asn	...	...	...	...	Arg
mOP-3	...	Arg	Asn	Asn	...	...	Val	...	Arg
DPP	Asn	...	Gln	...	Thr	...	Val	...	...
Vgl	...	Asn	Asn	Asp	...	...	Val	...	Arg
Vgr-1	...	...	Asn	...	...	...	...	...	...
CBMP-2A	...	Glu	Asn	Glu	Lys	...	Val	...	...
CBMP-2B	...	Glu	Tyr	Asp	Lys	...	Val	...	...
BMP3	...	Glu	Asn	Lys	...	...	Val	...	...
GDF-1	...	Asn	...	Asp	...	...	Val	...	Arg
60A	Leu	Asn	Asp	Glu	...	...	Asn	...	...
BMP5	...	...	...	...	...	...	...	...	...
BMP6	...	...	Asn	...	...	...	...	...	...

85

FIG. 1J

hOP-1	Lys	Tyr	Arg	Asn	Met	Val	Val	Arg	
mOP-1	...	...	...	...	...	...	...	...	
hOP-2	...	His	...	...	...	...	...	Lys	
mOP-2	...	His	...	...	...	...	...	Lys	
mOP-3	Arg	Glu	...	...	...	...	...	Gln	
DPP	Asn	...	Gln	Glu	...	Thr	...	Val	
Vgl	His	...	Glu	...	...	Ala	...	Asp	
Vgr-1	...	...	...	...	...	...	...	...	
CBMP-2A	Asn	...	Gln	Asp	...	...	...	Glu	
CBMP-2B	Asn	...	Gln	Glu	...	...	...	Glu	
BMP3	Val	...	Pro	...	...	Thr	...	Glu	
GDF-1	Gln	...	Glu	Asp	...	...	...	Asp	
60A	...	...	...	...	...	Ile	...	Lys	
BMP5	...	...	...	...	...	...	...	...	
BMP6	...	...	...	TrP	...	...	...	...	
	90					95			

FIG. 1K

hOP-1	Ala	Cys	Gly	Cys	His
mOP-1	...	...	...	...	...
hOP-2	...	...	...	...	...
mOP-2	...	...	...	...	...
mOP-3	...	...	...	...	...
DPP	Gly	...	...	...	Arg
Vgl	Glu	...	...	...	Arg
Vgr-1	...	...	...	...	...
CBMP-2A	Gly	...	...	...	Arg
CBMP-2B	Gly	...	...	...	Arg
BMP3	Ser	...	Ala	...	Arg
GDF-1	Glu	...	...	...	Arg
60A	Ser	...	...	...	...
BMP5	Ser	...	...	...	...
BMP6	...	...	...	...	...
			100		

**Between residues 56 and 57 of BMP3 is a Val residue; between residues 43 and 44 of GDF-1 lies the amino acid sequence Gly-Gly-Pro-Pro.

FIG. 1L

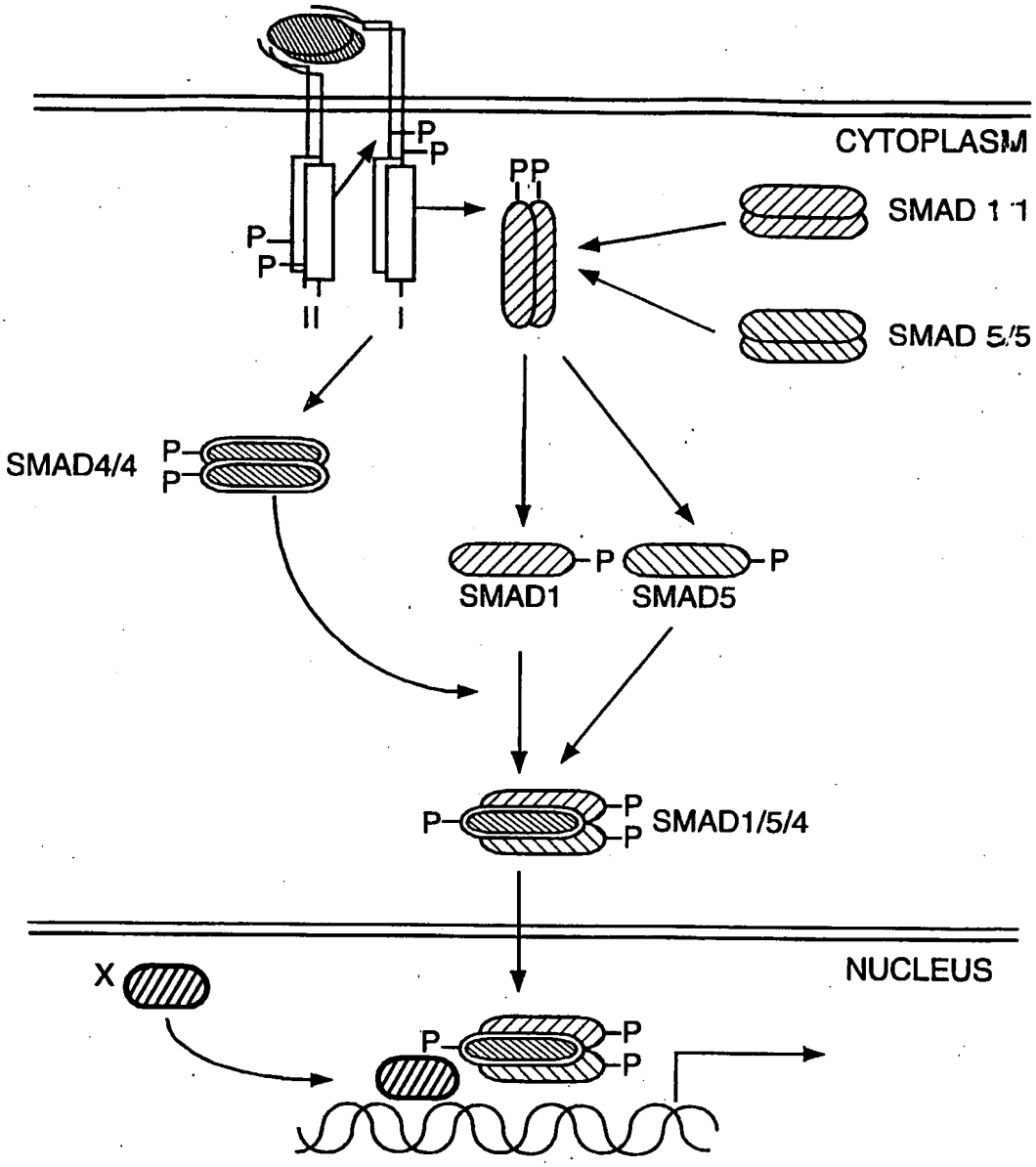
CYTOPLASM
P
P
P
P
II
I
PP
SMAD 1/1
SMAD 5/5
SMAD4/4
P
P
SMAD1
P
SMAD5
P
P
P
SMAD1/5/4
P
NUCLEUS
X
P
P
P

FIG. 2